SELF-PRIORITISATION IN SOCIAL COGNITION

By

Katherine Louise Jones

Department of Psychology

University of Warwick

This thesis is submitted for the degree of Doctor of Philosophy

September 2018

To Norris

CONTENTS

List of Tables

List of Figures

ACKNOWLEDGEMENTS

It feels like an understatement to say that I could not have done this alone. My heartfelt thanks go to everyone who has supported me, kept me going, and provided me with cups of tea. I appreciate you all.

First and foremost, I would like to express my deepest gratitude to my supervisor Derrick Watson. Thanks not only for your expert guidance, but also your unwavering support, respect, kindness, and confidence boosts. You truly are an exemplary researcher and mentor. It has been my privilege (and a lot of fun!) to work with you. I would also like to thank Melina Kunar, who saw potential in me and got me started on this endeavour.

Thanks also to my postgrad friends. Special mention goes to my 'co-dependents' Luke and Toria, the keepers of my sanity. It has been a joy. To Dan, Danni and Sophie: thanks for being so good to me, I will watch all your careers with great interest. To my patient and wonderful housemates; Emily, Matthew and most recently, Divya, thanks for putting up with me! Also, to the wonderful Tash, who has been an endless source of support.

On a more practical note, I am grateful to the ESRC who funded this research, and to all the participants who provided me with their valuable data. I also wish to thank the support staff who have saved the day on more than one occasion.

Finally, to my long-suffering but ever-loving family: you got me here. I am very fortunate to have the most supportive family imaginable; especially my parents Sally and Richard, and brother Alex. To my wonderful grandparents: you will forever be a source of inspiration. Most of all, I would like to thank my twin sister Anna, who understands more than anyone what this process has been like. Thanks for being there whenever I needed you.

DECLARATION

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author.

List of publications under revision:

Jones, K.L., Kunar, M.A., & Watson, D.G. (2018). I know what you're doing: Awareness of a o-actor's task set modulates own behaviour. Manuscript under revision with *Cognition*.

List of work presented at conferences and external seminars:

Jones, K.L., Kunar, M.A., & Watson, D.G. (2015) I know what you're doing! Awareness of other people's intentions interfere with cognitive task performance Poster presented at the European Conference of Visual Perception, Liverpool, UK.

Jones, K.L., Kunar, M.A., & Watson, D.G. (2016), What's Mine Isn't Yours: Self-Prioritization in Shared and Non-Shared Task Environments. Poster presented at the Psychonomic Society 57th Annual meeting, Boston USA

Jones, K.L, Kunar, M.A, & Watson, D. G (2017). Perceptual Salience of Self-Relevant Information in Shared Environments. *Journal of Vision*,*17*(10), 1288-1288. Poster presented at the Vision Science Society 17th Annual Meeting, St. Pete Beach, Florida USA

ABSTRACT

Systematic self-biases have been demonstrated across cognitive domains of perception, attention, memory, and decision making. However, it remains unclear how self-relevance is recognised and processed by the brain. This thesis examines the influence of social context on processes of self-prioritisation in social cognition.

Using a perceptual matching paradigm (Sui, He, & Humphreys, 2012), the first series of experiments demonstrated that novel shape stimuli associated with the self are prioritised for perceptual processing over novel shape stimuli associated with other individuals present within, and absent from, the task environment. No processing advantage was found for stimuli associated with present others over absent others. A second series of experiments demonstrated that self-associated stimuli were prioritised for processing over stimuli associated with both liked and disliked others. Moreover, the size of the self-association benefit was not influenced by the 'likability' of the other individuals. Therefore, neither the ability to 'tag' information to a physical body, nor the likability of social stimuli can fully account for self-prioritisation in perceptual matching. A final series of experiments introduced a novel paradigm to investigate the extent to which people prioritise own-task-relevant information over information relevant to a nearby actor's task. The findings indicated that participants co-represented a partner's stimulus-response rules even when interpersonal coordination was not required. Such co-representation occurred within a binary choice task in which direct stimulus response mappings were not possible. This suggests that co-representation is an unintended consequence of a shared social environment and that people are unable to fully prioritise their own task over that of a nearby actor.

Overall, the work extends and clarifies when and how self-biases are influenced by task demands, the context in which a task is performed, and the presence or absence of others within the environment.

1 SELF-BIASES IN COGNITION

Despite the remarkable processing power and complexity of the human brain it is severely capacity limited. At any one moment the sensory system receives more information than can ever be fully processed. Given this limitation, it is imperative that information most relevant to the self is prioritised for processing. Accordingly, notable self-biases have been found across a number of high- and low-level cognitive domains; such as attention (Keyes & Dlugokencka, 2014; Sui, Liu, Wang, & Han, 2009; Sui, Rotshtein, & Humphreys, 2013; Tacikowski & Nowicka, 2010), perception (Ma & Han, 2010; Sui, Chechlacz, & Humphreys, 2012; Sui & Humphreys, 2013), memory (M. A. Conway, 2005; Cunningham, Brebner, Quinn, & Turk, 2014; Klein & Kihlstrom, 1986b), and decision making (Beggan, 1992; Kahneman, Knetsch, & Thaler, 1990). Even though self-prioritisation effects are found to be ubiquitous in cognition, a fundamental question remains unanswered: how does the brain recognise and process self-relevance? In this chapter I provide an overview of research into self-biases across a number of cognitive domains, including investigations of how the self is represented within the brain.

1.1 Self-biases in Attention: The own-name effect

One of the earliest and well-known phenomenon in attention research is that of the 'Cocktail Party Effect' (Cherry, 1953). This effect describes the ability to tune in to a single conversation at a noisy party, while ignoring other conversations going on around you. This scenario is often used as an example of how individuals can selectively attend to one stream of information, while inhibiting other task-irrelevant streams. One of the notable features of this phenomenon is that certain high-priority stimuli, such as one's own name, are able to enter awareness even when presented within an unattended stream (Moray, 1959).

This "own name effect" was first demonstrated empirically by Moray (1959) using a dichotic listening task. Participants were asked to shadow (or verbally repeat) information as it was presented in one ear, while ignoring information presented to the other ear. Once the shadowing task had been completed, participants were asked if they recalled any details of the information presented to the unattended ear. Recall was typically poor, with participants only able to recount the physical properties of the information, rather than the semantic content, of the information (such as the accent, tone and gender of the speaker's voice). However, Moray found that participants had improved recall of information presented to the unattended ear if it followed the presentation of the participant's own name. Subsequent studies have also found that approximately 33 percent of participants are aware when their own name has been presented to the unattended ear (A. R. Conway, Cowan, & Bunting, 2001; Wood & Cowan, 1995). This influential finding had been taken to suggest the self-relevance of information is detected pre-attentively and can direct attention towards a previously unattended stream of information.

However, the idea that self-relevant information can be detected pre-attentively was at odds with early-selection theories of attention. For example, Broadbent's early-selection filter model proposed that information is filtered out *before* semantic processing, on the basis of the physical properties of the information, such as the tone

of a speaker's voice, or which ear the sound was being played to (Broadbent, 1958). Treisman (1960, 1964a) later revised Broadbent's filter theory to suggest that the filter *attenuates* unattended information, rather than filtering it out completely. This attenuation makes it more difficult, but not impossible, to extract meaningful content from 'unattended' streams. Importantly, Treisman further suggested that words with subjective importance, such as one's own name, have lower activation thresholds for recognition than other types of information. Therefore it is easier for self-relevant stimuli, such as one's own name, to enter awareness even when presented within an attenuated stream (Treisman, 1964b).

Alternative theories propose that self-relevance cannot be detected pre-attentively. Instead, it may be that participants occasionally monitor the unattended stream, or that attention unintentionally 'slips' from one stream to the other (Gaspelin, Ruthruff, & Jung, 2014; Lachter, Forster, & Ruthruff, 2004). The enhanced reporting of own-name stimuli compared to other information presented to the unattended ear may reflect a bias for reporting self-relevant information, rather an attentional bias (Humphreys & Sui, 2016).

In the six decades since Moray's influential experiment, a number of studies have found evidence to suggest that own-name stimuli do not, in fact, capture attention. For instance, Harris, Pashler and Coburn (2004) conducted a series of visual search tasks and found that although participants were faster to detect own-name targets than other targets, the search slopes were not shallow enough to suggest that the name 'pops out' at the participant. Furthermore, own-name stimuli were not found to act as particularly potent distractors compared to other-name stimuli. Taken together, these findings suggest that the prioritised processing of self-relevant stimuli occurs only once attention has been allocated to the stimuli.

Gronau, Cohen, and Ben-Shakha (2003) conducted a study in which they directly investigated whether self-relevant information can capture attention when presented outside of the focus of attention. Using a Stroop-like task, participants were asked to respond to the colour of a stimulus presented at the centre of a display, while

ignoring word stimuli presented either at the periphery (i.e., outside of the focus of attention), or also at the centre of the display (i.e., within the focus of attention). When the word was presented at the centre of the display, the naming of the colour stimulus was most disrupted if the word was the participant's own name than another person's name. However, this behavioural effect did not occur when the name stimuli were presented at the periphery. The data from this study suggest that one's own name receives enhanced processing *once* it has been attended to. However, own-name stimuli presented at the periphery were unable to capture attention. Of note, the researchers also measured the galvanic skin responses of the participants throughout the task. The data revealed a heightened response to own-name stimuli compared to other stimuli, even when they were presented at the periphery. This finding suggests that even though attention needs to be allocated to an own-name stimulus for it to break into awareness, there may be some implicit recognition of self-relevance.

Although own-name stimuli may not capture attention, there is evidence to suggest that, under certain conditions, one's own name can act as a powerful cue for attention. Alexopoulos, Muller, Ric and Marendaz (2012) used names as cues for the location of a target prior to the presentation of a visual search display. Own-name cues were found to facilitate search to a greater degree than other-name cues. This cueing effect occurred even when the names were masked so that participants were unable to judge whether the name was theirs or not. Furthermore, the authors conducted an 'anti-saccade' version of the cueing task, in which the cue was presented at the opposite side of a central fixation from the target. Therefore, in this version of the task, participants were required to saccade *away* from the cue, rather than towards it. The data revealed that the presentation of an own-name cue disrupted performance to a greater extent than other-name cues, indicating that it is particularly difficult to disengage attention from own-name stimuli. The authors suggested that the own-name cueing effect is an indication that personal names can capture attention. However, it has been pointed out that as only a single cue appeared on the display, participants were likely to have adopted a distributed mode of attention when performing the search task (Humphreys

& Sui, 2016). Therefore, it may be that the own-name stimuli are only a powerful cue for attention when there are sufficient attentional resources.

However, a number of studies have found that own-name effects often persist under conditions in which attentional resources are limited (Arnell, Shapiro, & Sorensen, 1999; A. R. Conway et al., 2001; Wood & Cowan, 1995). For example, own-name stimuli have been found to survive the "attentional blink" (Shapiro, Caldwell, & Sorensen, 1997). The attentional blink describes the impaired ability to detect the second of two targets presented in close temporal succession (Raymond, Shapiro, & Arnell, 1992). The paradigm employed to investigate the attentional blink is known as rapid serial visual presentation (RSVP). Stimuli such as digits, letters, or images are presented in rapid succession at a single location at rates between 6 – 20 items per second. The attentional blink was originally thought to reflect the temporary unavailability of attentional resources for the second target while the first is being processed. However, there is an ongoing debate as to whether the attentional blink reflects an absolute capacity limitation, or the suboptimal allocation of resources. For instance, Olivers and Nieuwenhuis (2006) found that increasing task load by asking participants to perform a concurrent memory task, *reduced* the likelihood of an attentional blink. This finding was taken to suggest that divided attention facilitates performance because resources are more evenly distributed. In contrast, Giesbrecht, Sy and Lewis (2009) found that when the perceptual load of the first target was increased, personal names no longer survived the attentional blink. Therefore, in this study, increased perceptual load was found to negatively impact the selectivity of attention over time.

What is apparent, is that for own-name stimuli to survive the attentional blink the self-relevance of the name must first be processed and detected, at least on an implicit level. Studies measuring event related potentials (ERPs), which indicate with high temporal resolution the brain activation associated with one or more mental operations, have found that hearing and reading one's name is associated with an enhanced P300 component (Tacikowski, Cygan, & Nowicka, 2014; Tacikowski & Nowicka, 2010). Although this component has been associated with processes of

attention, it is thought to occur too late to reflect the early capture of attention (Dien, Spencer, & Donchin, 2004). Interestingly, the attentional blink has been associated with the complete suppression of the P300 (Sergent, Baillet, & Dehaene, 2005). It is not clear how these two findings co-exist if the detection of self-relevance is to protect information from the attentional blink.

Summary: A large body of research has suggested that own-name stimuli are subject to prioritised attentional processing, although it is unclear whether they can capture attention. Converging evidence from behavioural and neurophysiological studies suggest that self-biases are dependent on the allocation of attention rather than occurring pre-attentively. Nevertheless, own-name stimuli have been found to be more likely to enter awareness than other forms of stimuli when attentional resources are limited. It therefore appears that the self-relevance of information can bias the processes of attentional allocation.

1.2 Self-biases in Memory

It has long been established that people have superior memory for information encoded in reference to the self, compared to information with no, or weak, self-relevance – a phenomenon known as the self-reference effect (Klein & Kihlstrom, 1986a; Rogers, Kuiper, & Kirker, 1977; Symons & Johnson, 1997). The discovery of the self-reference effect was motivated by the influential depth-of-processing model of cognition (Craik & Lockhart, 1972; Craik & Tulving, 1975), which proposed that deep semantic encoding of information results in more robust and enduring memory traces.

Evidence for the depth-of-processing effect originated from experiments demonstrating that semantic encoding tasks (e.g., "Does the word mean the same as [insert word]"?) results in better recall of word stimuli that phonemic encoding (e.g., "Does the word rhyme with [insert word]"?) or structural encoding (e.g., "Does the word have upper case letters"?) tasks (Craik & Tulving, 1975). Rogers et al. (1977) expanded on this paradigm by adding a self-reference encoding task (e.g., "Does [insert

word] word describe you"?). Memory performance for the words encoded in reference to the self was superior to all other conditions, including the semantic encoding condition. This self-reference effect has been replicated multiple times over the last four decades (Bentley, Greenaway, & Haslam, 2017; Cunningham et al., 2014; Klein, 2012), and a large-scale meta-analysis has supported the existence of a substantive mnemonic benefit of self-reference (Symons & Johnson, 1997).

Cunningham, Turk, Macdonald and Macrae (2006) found that the self-reference effect not only occurs when using word stimuli, but also arbitrary objects only recently associated with the self. In their novel paradigm, participants were asked to sort picture cards depicting grocery items into one of two shopping baskets, one that belonged to themselves and one that belonged to another participant. After completing the sorting task, participants were given a surprise recognition test in which they were required to identify previously seen shopping items from amongst previously unseen filler items. Participants were more likely to recognise items they had sorted in to their own basket than items they had sorted in to the other person's basket, independent of whether the participant owned the object in real life. The results of this experiment demonstrate that simply 'tagging' novel objects with self-relevance increases their chance of being remembered.

More recent research has suggested that people not only have better general recall for self-referent information, but that they incorporate more perceptual and contextual details within those memories (Sui & Humphreys, 2015c). For example, Leshikar, Dulas and Duarte (2015) asked younger and older adults to judge positive and negative adjectives either for their self-descriptiveness or for semantic meaning. There was an overall memory advantage for adjectives encoded in reference to the self over words encoded based on their semantic meaning. Participants were also more likely to remember perceptual and contextual details about the words they judged for self-descriptiveness; such as whether the words had been spoken by a male or female voice; whether the words were presented in a red or green; the temporal order in which the words were presented; and the emotions that had been evoked by the words. This finding that people integrate more perceptual and contextual detail within memories

encoded in reference to the self held for both the younger and older adults. Therefore, it appears that self-reference promotes the binding of different forms of information within memory, such as verbal and perceptual details.

A study conducted with young children aged between 4 and 6 years old also found that children had superior recall and source memory for items encoded in reference to the self (Cunningham et al., 2014). The children were presented with a series of images depicting either themselves or a same-aged child and a unique object was presented next to each image. The children were asked to judge whether they thought the child in the image (self or other) would like the presented object. In a subsequent recognition test of the objects presented with the images, the children were more likely to recognise the objects presented alongside an image of themselves than an image of another child. Furthermore, the children demonstrated better memory for the image the object had been presented with, and the location of the object relative to the image, if the image was of themselves.

A number of theories have been proposed to explain why encoding information in relation to the self facilitates memory. For one, it is known that elaboration improves memory (Craik & Tulving, 1975; Eysenck & Eysenck, 1979), and it is possible that self-reference promotes increased elaboration. For instance, when asked to judge whether an adjective describes the self (e.g., are you tidy?), a participant might think of a number of different scenarios in order to determine if the adjective would be fitting. This process of elaboration then acts to strengthen the memory trace (Klein & Kihlstrom, 1986a). Another proposal is that the participant's self-construct connects and organises a large amount of pre-existing information about the self, and this organisational structure facilitates the retrieval of information from memory (Reuther & Chakravarthi, 2017).

Of note, a neuropsychological study has found evidence to suggest that self-referential processing is distinct from elaborative processing (Sui & Humphreys, 2013a). A patient, GA, who had both sematic impairment and severe amnesia, was shown a number of images that were either assigned to them or to the experimenter. In a subsequent memory test GA had better memory for the items that had been assigned to

them. The size of this self-reference effect was similar to that of healthy age-matched controls. In a second task, GA was asked to either semantically classify each item (i.e., living or non-living) or judge the item based on its physical properties (e.g., large or small). The results revealed that GA was no better at recognising items that had been processed at a deep semantic level than items judged based on their surface properties. This is in contrast with the healthy controls, who showed superior recognition and source memory for the items that had undergone semantic processing. Therefore, although GA demonstrated a robust self-prioritisation effect, they demonstrated no benefit of semantic encoding.

Self-reference effects may be less a product of elaborative encoding and more a product of the organisation of self-relevant information. For instance, if an individual judges a trait adjective for self-descriptiveness they may then encode the words within categories of whether they are good or bad descriptors of the self (Reuther & Chakravarthi, 2017). These categories can then be used to facilitate recall (Klein & Kihlstrom, 1986a). Furthermore, by incorporating information within a well-established network of self-relevant information (i.e., one's self-construct) new information becomes connected with large body of pre-existing information, and these connections likely result in improved memory (Klein & Kihlstrom, 1986b; Klein & Loftus, 1988).

Summary: There is much evidence to suggest that information encoded in reference to the self is better remembered than information encoded in reference to others. This self-reference effect may be due to enhanced elaborative or organisational processes. Importantly, it has also been found that people have superior source memory for self-relevant information, and that they incorporate more perceptual information within the memories of self-relevant items. The idea that self-relevance acts to integrate processes of perception and memory has become a recent topic of interest (Sui, 2016; Sui & Humphreys, 2015c). This topic will be more fully addressed in following sections.

1.3 Self-biases in Face Perception

Humans show distinct behavioural responses to their own face than to the faces of other individuals (Keyes & Dlugokencka, 2014; Ma & Han, 2010; Sui & Han, 2007). For example, Tong and Nakayama (1999) found that when performing a visual search task, participants were more efficient at detecting an own-face target than a stranger-face target from within a display of distractors. This benefit was found to be view invariant; that is, it did not matter whether the target stimuli were presented face-forward or at profile, upright or upside-down. Tong and Nakayama suggested that the own-face advantage reflected the fact that one's own face is a highly familiar and over-learned stimulus.

Recent research has also found evidence to suggest that one's own face is a more effective cue for attention than the faces of other people. Liu, He, Rotsthein and Sui (2016) asked participants to judge the orientation of a letter stimulus presented either in the left or right periphery of the display. First, an image of a front-facing face was presented in the centre of the display. The image then dynamically changed so that the face was orientated to either the left or the right side of the display, cueing the location of the target. Then, the target appeared with a varying stimulus onset asynchrony. If the stimulus onset asynchrony was long (150ms), a cueing effect was found irrespective of whether the face was the participant's own face, or the face of their best friend. However, if it was short (50ms), the self-cue exclusively facilitated attention to the cued target. The results of this experiment suggest that own-face stimuli are processed faster and are therefore a more efficient cue for orienting attention.

This behavioural data was supported by ERP findings. The self-face cues elicited greater amplitudes of the N1 component over the left central-parietal region, a relatively early component associated with the neural processing of faces (Itier & Taylor, 2004), compared to the friend-face. The magnitude of the cueing effect in the N1 component predicted a reduction of amplitude of the P3 component over the right frontal-central-parietal cortex. This relationship between the two components was thought to reflect

that the self-face cuing benefit originated from the enhanced early sensory processing of the self-face, which then reduced uncertainty of the location of the target, reducing the P300 component.

There is now converging evidence that own-face processing is associated with relatively early components of the visual evoked response. For example, Sui, Zhu and Han (2006) asked participants to judge the head orientation of self-faces and faces of familiar or unfamiliar others. Relative to other-face stimuli, the processing of own-face stimuli was associated with increased positivity over frontal-central regions within 220-270 ms from stimulus onset (the anterior N2). Furthermore, this effect on N2 magnitude was also found when face-stimuli were used as distractors and participants were instead asked to make judgements about the central fixation cross. The authors suggested that self-face processing occurs automatically once face-structure encoding has occurred and is independent of task-relevance.

Similarly, Keyes, Brady, Reilly, and Foxe (2010) asked participants to monitor a series of images of their own face and faces of friends and strangers. Differences in ERP waveforms were found to occur relatively early, such as increased amplitudes of the N170 and the VPP components over posterior and fronto-central sites for self-faces relative to both friend and stranger faces. This self-effect was found to occur at approximately 250ms from stimulus onset. In comparison, differences in responses to friend and stranger faces were only marked after 250ms. Taken together, the results of these studies suggest that own-face stimuli evoke a rapid response thought to reflect either enhanced early perceptual processing, or the prioritised allocation of attention to stimuli. Furthermore, this response occurs even when the presented face is not task related.

Other studies have also found evidence that self-face effects can occur even when the participants are not consciously aware of the presence of the face. For instance, Geng, Zhang, Li, Tao and Xu (2012) used dynamic noise patterns to supress awareness of face stimuli. The presentation of a supressed face-stimulus was followed by the presentation of word stimulus, and participants were asked to judge if the word

had a positive or negative valence. Participants with high self-esteem were found to respond faster to positively-valenced words if they followed the presentation of a supressed self-face than a famous face. This suggests that the self-face stimuli were faster to gain sufficient dominance over the suppression in order to influence responding.

However, other studies suggest that own-face effects do not occur pre-attentively. Keyes and Dlugokencka (2014) investigated whether self- friend- and stranger- faces provide differential levels of distraction when processing self, friend and stranger names. The word stimuli were either superimposed over the face stimuli, so that the face was within the focus of attention, or were flanked by the face stimuli presented at the periphery. When the face distractors were presented at the focus of attention, self and friend images facilitated self and friend naming, respectively. However, when the face distractors were presented at the periphery no such facilitation effect was found. Therefore, the finding of this study suggest that own-face effects only occur if the face stimuli are within the focus of attention.

It is not fully clear what cognitive mechanisms drive the own-face advantage. Ma and Han (2010) investigated if the perception of self-face stimuli implicitly activate positive attributes relating to one's self-concept, and whether this positive association in turn facilitates behavioural responses to one's own face. Prior to judging the orientation of face stimuli, participants were either primed to think negatively or positively about themselves. This was done by asking them to judge either positive or negative trait adjective for self-descriptiveness. When participants were primed to have a positive association with themselves, they were faster to judge the orientation of self-face stimuli relative to other face stimuli. However, when participants were primed to think negatively about themselves this self-face advantage disappeared. The authors proposed that an implicit positive association with the self may be driving self-biases in face perception.

It is possible that the positive evaluations of face stimuli are confounded by self-face similarity. Social research has found that most people are inclined to believe that

people who are similar in physical appearance to themselves will be more trustworthy and likable (DeBruine, Jones, Little, & Perrett, 2008). In a recent study conducted by Finke, Larra, Merz, and Schächinger (2017), participants were shown images of morphed faces that were manipulated to look like their own or an unknown face. Self-resemblance was found to induce deeper processing of facial affect, particularly of happy facial expressions. A second study using famous faces and unknown faces showed a similar effect of familiarity, however unlike the previous experiment participants were no faster to respond to familiar faces than unknown faces. The results of this study suggest that not only self-faces, but faces similar to one's own, are processed differently from face stimuli unrelated to the self. Furthermore, there is evidence of an early interplay between the recognition of facial identity and the interpretation of emotional expression during face processing. The study also found both similarities and dissociations between the effects of self-relatedness and familiarity.

Sui and Humphreys (2013b) proposed that the self-advantage in face processing reflects the fact that people use the self as an anchor to make self-other discriminations during the categorisation of faces. The authors found that the distribution of response times to self-faces did not alter when participants categorised faces based on belonging to: (i) the self OR the friend/stranger, or (ii) the stranger OR the self/friend. However, there were shifts in the response distributions for the friend and stranger faces. This pattern of results was taken to suggest that there exists a qualitative distinction in the processing of self-faces and other faces. This distinction is unlikely to be familiarity, as the friend's face is likely to be highly familiar to the participant, whereas the stranger face is highly unfamiliar. Rather than there being a linear relationship between responses to the self-, friend- and stranger-faces, the authors proposed that self-faces belonged to a distinct category.

Summary: People are not only faster to detect their own face relative to the faces of familiar and unfamiliar others, they are also more efficient at making perceptual judgements about own-face stimuli. There is mixed evidence as to whether this self-face advantage occurs automatically, however ERP studies suggest that the self-face bias occurs early in the processing stream, possibly relating to the allocation of attention or

initial perceptual processing. Possible factors driving the self-face advantage include the familiarity, implicit positive value, and category distinctiveness of 'self' relative to 'other'.

1.4 Is 'The Self' Special?

So far, research demonstrating robust self-biases across cognitive domains of attention, memory and face perception have been discussed. The ubiquity of self-biases in cognition prompt the question of whether the self has a unique status in cognition. However, in order to answer this question, it must first be determined whether more generic cognitive biases are responsible for driving these effects. For instance, it is possible that the enhanced familiarity, the inherent emotional valence, or the intrinsic reward value of self-related stimuli are responsible for driving self-biases. Henceforth, the term 'general cognitive bias' will be used as an umbrella term to encompass all cognitive biases that are *not* specific to the processing of self-relevant information but may contributen to a self-association advantage. The following sections will outline research that has attempted to disassociate self-biases from other potentially confounding general cognitive biases.

1.5 Disassociating the effects of self-relevance and familiarity

The majority of studies investigating self-biases in cognition often use over-learned self-related stimuli, such as the participant's own name or face. As a result, it is difficult to determine whether the prioritised processing of such stimuli are due to the association with the self, or because they are highly familiar. Sui, He and Humphreys (2012) devised a now highly influential paradigm to get around the confound of stimuli familiarity. In their novel perceptual matching task, participants associated arbitrary shape stimuli (e.g., a triangle, square, and circle) with the self, a familiar other, and an

unfamiliar other. For example, participants could be told the following "you are represented by a triangle, your best friend is represented by a circle, and a stranger is represented by a square". Once these shape-person associations had been encoded in memory, participants were then presented with a random combination of a shape with a social label referring to the self, the familiar-other or the unfamiliar-other. The participants were required to make speeded judgements as to whether the presented shape and label stimuli conformed to the just-learned associations (e.g., triangle-self), or whether they were a recombination of a shape with a non-matching label (e.g., triangle-stranger).

The results of this initial experiment were striking, participants were considerably more efficient (faster and more accurate) to respond to matched shape-label pairings if the shape was associated with the self than if it was associated with either the friend or the stranger. Note that the shape stimuli were all essentially 'neutral'; in that they had no pre-learned social associations (but see D. G. Watson, Blagrove, Evans, & Moore, 2012), the shape stimuli were equally familiar, and the assignment of social relevance was counterbalanced across participants. Accordingly, any difference in responses as a function of shape are likely due to newly learned social associations, rather than differences in familiarity.

A second experiment was conducted in which the friend association was replaced with a mother association. It can be argued that the concept of one's mother is as familiar as one's self-concept, and may even have been established earlier in development (Damon & Hart, 1982; Montemayor & Eisen, 1977). Furthermore, research has indicated that self-representation has strong associations with mother associations, especially in some cultures (Han & Northoff, 2009; Zhu & Han, 2008; Zhu, Zhang, Fan, & Han, 2007). However, Sui et al. (2012) found a reliable advantage of self-association over mother-association. Therefore, irrespective of purported conceptual overlap between representations of oneself and one's mother, there remained a clear advantage of self-association over mother-association during perceptual matching.

A series of control experiments indicated that the self-prioritisation effect was not a result of the concreteness, frequency, or length of the different social labels. Also, a self-association benefit was found for different self-labels, including 'you', 'yourself' and the participants own name. Importantly, this advantage occurred when the label was presented before the shape stimulus, so that the participant responded primarily to the presentation of the novel shape, minimising the effects of label familiarity. Furthermore, the self-association advantage occurred even when the shape stimuli were degraded by reducing the contrast of the shape stimulus relative to the background. Whereas visually degrading shape stimuli disrupted responses for the familiar and unfamiliar other associations, there was no such effect of stimulus degradation of the self-associated shape. This latter result is indicative of self-association influencing the early stages of perceptual processing.

Sui, Sun, Peng and Humphreys (2014) further investigated the role of stimulus familiarity on perceptual matching performance by manipulating the frequency with which the stimuli associated with the self, mother, and stranger, appeared. The authors were interested in whether having a strong expectancy for the mother- or stranger-shape matched pairings would override the bias for self-shape matched pairings. Perceptual matching performance was compared between blocks in which the three matched pairings occurred at an equal frequency, and blocks in which the three matched pairs occurred at a ratio of 1:3:3 (with either the self-, friend- or stranger- matched pairings occurring at the lower frequency).

The authors reported a number of interesting findings. First, for responses to the self-matched pairs, there was no cost of low frequency relative to baseline (where all matched trials occurred equally often). In contrast, for the mother- and stranger-matched pairings there was a cost of being a low-probability occurrence (i.e., slower RTs). Second, the magnitude of the self-association benefit was greater when the self-matched pairings were a high-frequency occurrence. Thus, holding an expectancy for the self-matched pairs had a benefit on performance. Finally, performance for the mother- and stranger- stimuli improved when they both occurred at an equal-high probability within the same block. In contrast, when either the mother- or stranger-

matched pairs occurred at an equal-high probability as the self-matched pairs, there was only a benefit of self-association. Therefore, expectancies held for the self-stimuli appeared to override expectancies for the other stimuli.

The results of this study suggest that biases for self-related stimuli occur automatically, irrespective of the probability of occurrence, and are difficult to override. In contrast, expectancies for stimuli associated with other individuals can bias performance, but this bias can be overridden by an expectancy for self-related stimuli. This is an indication that the effects of self-association in perceptual matching tasks operate in a relatively exclusive manner and cannot be fully accounted for by stimulus familiarity.

Summary: Using a simple yet elegant perceptual matching paradigm, Sui, He and Humphreys (2012) found that participants were significantly faster to match novel stimuli to the self than to familiar or unfamiliar others. This finding suggests that self-association can bias processes of attention and perception, and that these effects are independent from those of stimulus familiarity. Furthermore, the self-bias was found to be robust, surviving even when the presentation of self-stimuli was a low-probability event. Overall, the findings from perceptual matching studies suggest that the self-bias is not equivalent to a familiarity bias.

1.6 Disassociating the effects of self-relevance and reward

Along with the effects of stimulus self-relevance, there is evidence to suggest that the reward value of stimuli can modulate processes of perception and attention (Hickey, Chelazzi, & Theeuwes, 2011; Peck, Jangraw, Suzuki, Efem, & Gottlieb, 2009). For example, Anderson and Yantis (2013) found that stimuli associated with reward involuntarily capture attention when presented as distractors in a visual search task. In contrast, Hickey et al. (2011) found that reward improved target selection during visual search, but had no effect on the suppression of distractors.

It is therefore possible that self-biases in cognition reflect the reward value of information associated with the self. Evidence from neuroscientific studies have provided evidence of some overlap between the neural substrates of self and reward (de Greck et al., 2008; Enzi, de Greck, Prösch, Tempelmann, & Northoff, 2009; Enzi et al., 2009; Northoff & Hayes, 2011). For example, Enzi et al. (2009) conducted functional magnetic resonance imaging (fMRI) on participants while they performed a reward gambling task and a personal relevance attribution task. The two tasks were found to activate overlapping areas of the brain, including the ventromedial-prefrontal cortex (vmPFC), ventral striatum and ventral tegmentum.

A number of behavioural studies have also been conducted to investigate whether self-bias is equivalent to a reward bias. For example, Sui et al. (2012) found that assigning a reward value to novel visual stimuli can mimic the effect of self-relevance during a perceptual matching task. Participants were asked to associate shape stimuli with either a high reward (£9), medium reward (£3) or low reward (£1) values. The participants then carried out a perceptual matching task in which they had to determine whether a shape stimulus had been presented with a label representing the appropriate reward value. On each trial, participants received a monetary reward proportionate to the value presented. In much the same way as responses to self-relevant stimuli were faster than responses to other-relevant stimuli, responses to the high-reward stimuli were faster than responses to low reward stimuli. Furthermore, both self-relevant and high-reward stimuli were found to be affected less by stimulus degradation than stimuli associated with a familiar/unfamiliar other or low reward. Although Sui et al. (2012) found that both reward and self-association modulate perceptual processing in a qualitatively similar manner, it was still unclear if the two effects have a common origin.

More recently, Sui and Humphreys(2015d) directly investigated the interaction between self-relevance and reward under perceptual matching conditions by rewarding participants with different monetary values when they correctly responded to self-, friend-, and stranger-matched shape-label pairs. Self-associated stimuli were always associated with no monetary reward (across all blocks of trials). However, on some blocks the friend was associated with a high reward and the stranger with low reward,

and on the other blocks the reverse was true. Despite the fact that participants received no reward when responding to self-associated stimuli, there was a notable benefit of self-association over both high and low reward stimuli. There was also an advantage for high-reward over low-reward across all blocks (irrespective of whether the high reward was associated with the friend or the stranger). Despite the fact that self-association and high-reward biased performance, no correlation between the two biases was found at the level of the individual. This lack of correlation suggests that advantages for self-association and high reward have distinct origins.

Furthermore, Sui and Humphreys (2015d) found that the magnitude of the self-association advantage, and the independence of self-bias and reward, varied dependant on how participants rated their perceived personal distance between the themselves and the stranger. For individuals who perceived themselves to have a close personal distance to the stranger, there was a smaller overall self-association advantage, and sensitivity to reward was found to modulate the relationship between self- and high-reward associations. In contrast, for individuals who indicated a larger personal distance between themselves and the stranger, sensitivity to reward did not modulate the magnitude of the self-association benefit. The authors interpreted this pattern of results to suggest that there was a common subjective value system for self-relevance and reward, but only for individuals with a close personal distance from others.

In a different study, Sui and Humphreys (2015a) examined the relative effects of reward and self-relevance on the integration of perceptual stimuli. To do so, they measured redundancy gains, which is the performance benefit found when two exemplars of a target are present, compared to just one (J. Miller, 1982; Raab, 1962; Wenger & Townsend, 2006). The magnitude of the redundancy gain is thought to reflect the extent to which target stimuli integrate together (Yankouskaya, Booth, & Humphreys, 2012). Sui and Humphreys (2015a) applied this logic to perceptual matching performance. In one experiment participants associated two shapes with the self and two shapes with a named best friend. In another experiment participants associated two shapes with a high reward value and two shapes with a low reward value. Once the associations had been encoded participants were presented with either single-

item or two-item shape displays. On the two-item display the shapes could either be identical, or were two different shapes associated with the same person or reward value. Participants were asked to identify whether the display contained shape(s) associated with the self versus the friend, or with high value versus low value.

The data revealed that the magnitude of the self-association bias increased significantly when more than one self-associated shape was presented on the display. Importantly, this redundancy gain occurred irrespective of whether the two self-associated shapes were perceptually the same, or were two different shapes associated with the self. There was also an overall benefit of high-reward over low-reward, although there was no redundancy gain if the two exemplars were the same shape (i.e., both perceptually and conceptually redundant). However, there was a redundancy gain when the two high-reward shapes were perceptually different. This pattern of results suggests that reward only modulated high-level conceptual processes, rather than low level perceptual integration. However, self-association modulated stimulus integration at both conceptual and perceptual levels. Therefore, the authors concluded that the effects of self-relevance on the perceptual and conceptual processing of shape stimuli are not the same as the effects of reward.

Further evidence of a dissociation between the effects of self-relevance and reward on perceptual processing come from neuropsychopathology case studies. Sui, Enock, Ralph and Humphreys (2015) found that a patient, RR, who had damage to the posterior superior temporal sulcus (pSTS; Sui et al., 2013) resulting in a hyper-self bias, also demonstrated a hypo reward-bias. These opposite biases for self and reward suggest they operate on two distinct value systems. However, a second patient, SC, who had damage to the vmPFC (thought to be the seat of self-representation; Humphreys & Sui, 2016) demonstrated no self-association benefit during perceptual matching *and* was less affected by reward than the healthy controls. The co-occurrence of a hypo self-bias and a hypo reward-bias as a result of a specific brain injury suggests that they could involve common processes. Nevertheless, the important behavioural finding is that self- and reward-biases can operate separately, and therefore are unlikely to have fully overlapping neural substrates.

Summary: Neuroimaging and behavioural data suggest that self-biases in attention cannot be fully accounted for simply as a reward-bias. Although the effects of self-relevance and reward can be qualitatively similar in some respects, notable difference have been found.

1.7 Disassociating the effects of self-relevance and positive self-evaluation

It has long been assumed that much of human behaviour is motivated by a desire for positive self-regard (James, 1950). For example, people have a tendency to describe themselves using more positive personality adjectives than they use to describe other people (Alicke & Sedikides, 2009; Kwan et al., 2007). People are also faster to process and better at recalling positive information about the self compared to negative information (Kuiper & MacDonald, 1982). Tests of implicit attitudes have found that most people hold an overall positive self-evaluation (Greenwald & Banaji, 1995; Koole, Dijksterhuis, & Van Knippenberg, 2001). It is possible, therefore, that people demonstrate enhanced responding to self-relevant stimuli because of an inherent positive association with the self.

Indeed, there is now evidence to suggest that positive self-regard may play a role in driving cognitive self-biases. For example, a recent ERP study found that positive self-evaluation influences the processing of incoming self-relevant information (Fields & Kuperberg, 2015). Participants were asked to read sentences describing scenarios that involved either themselves or another person. Positively valenced words processed within a self-relevant sentence resulted in a reduced N400 component relative to when the same word was processed within the context of a scenario involving another individual. This effect was thought to reflect an expectancy for positive information when reading about the self. Furthermore, Ma and Han (2010) found that positive self-evaluation plays an important role in the self-advantage for face perception. When

participants were primed to evaluate themselves negatively, by asking them to judge how well unfavourable trait adjectives describes themselves, participants were no better at making perceptual judgements about their own face than about other faces. This finding was taken to suggest that the activation of positive self-attributes when processing self-related stimuli in turn facilitates processing.

In a recent study, Stolte, Humphreys, Yankouskaya and Sui (2017) attempted to disassociate the effects of self-relevance and positive emotion using an adapted version of the Sui et al. (2012) perceptual matching paradigm. Participants performed perceptual matching tasks in which they either associated geometric shapes with social labels varying in self-relevance (i.e., self vs. friend- and stranger- associations) or with drawings of faces with varying emotional expressions (i.e., happy, neutral and sad). The data revealed a self-association bias comparable in magnitude to that found in previous experiments (Sui et al., 2012). There was also a reliable advantage for the faces with happy expressions over faces with neutral or sad expressions.

The magnitude of the positivity bias did not vary significantly if the facial expression was 'happy' or 'very happy'. This finding suggests that valence rather than the intensity of the emotion was driving the positivity advantage. At the individual level, both the self-bias and the positivity bias reflected trait-like behaviours, as they were found to be stable over the course of several weeks. Although the effects of self-relevance and positivity were found to be qualitatively similar, they were not found to correlate at the level of the individual. This lack of correlation indicates that the two biases have distinct underlying mechanisms.

Several neuroimaging studies have also demonstrated a dissociation between emotional and self-referential processing. For example, Moran, Macrae, Heatherton, Wyland and Kelley (2006) asked participants to judge favourable and unfavourable trait adjectives (e.g., 'honest' or 'lazy') for personal relevance, while undergoing functional magnetic resonance imaging (fMRI). Activation within the medial prefrontal cortex was associated with the self-descriptiveness of the trait adjectives, but not the emotional valence. Instead, the ventral anterior cingulate was associated with discriminating the

valence of the trait adjectives. Northoff et al. (2009) found evidence of a dissociation between self-relatedness and emotion within the same brain regions, suggesting that the processes are linked but not equivalent.

Summary: People typically hold both implicit and explicit positive associations with the self. It is therefore possible that self-biases in cognition are simply a manifestation of a positivity bias. However, both neuroimaging and behavioural studies have been able to disassociate the effect of self-association from positive emotion, suggesting that they have distinct origins.

1.8 Neural underpinning of self-bias

The above sections have presented evidence that self-biases are pervasive across processes of perception, attention, and memory. Furthermore, these purported self-biases are unlikely to be driven by the enhanced familiarity of self-related information, rather than the inherent reward value or positive emotion associated with the self. This goes to suggest that there may be some form of unique mechanism involved in the processing of self-relevant information. A number of studies have been conducted to determine the neural underpinnings of self-bias.

Sui et al. (2013) recorded fMRI data as participants matched novel shape stimuli to labels referring to the self, familiar others, and unfamiliar others (Sui et al., 2012). When participants responded to self-associated shape-label pairings there was enhanced activity in the area of the brain associated with self-representation: the ventral medial prefrontal cortex (vmPFC). As well as enhanced activation in the left posterior superior temporal sulcus (LpSTS), an area of the brain associated with social attention. The authors found that enhanced responses to the self-associated stimuli were associated with a strong top-down connection between activation in the vmPFC and the LpSTS. This has been taken as evidence that the processing of self-associated stimuli rapidly activates cognitive self-representations, which in turn primes the attentional system to

prioritise the processing of stimuli associated with the self. Therefore, the data indicates that the assignment of personal significance to arbitrary stimuli biases perceptual matching by strengthening the coupling of brain regions involved in internal self-representation with areas of the brain modulating attentional responses to social stimuli.

Importantly, responses to shape stimuli associated with other individuals (friend and stranger) were associated with enhanced activations within the fronto-parietal network, which has been implicated in processes of attentional control (Fassbender et al., 2006; Wang et al., 2010). Furthermore, activation of the fronto-parietal network was inversely correlated with that of the vmPFC and the LpSTS. Sui et al. (2013) suggested that this neural circuitry fits the idea that: (i) responses to the self-associated stimuli are independent of the attentional control network, and (ii) the attentional control network is being recruited for the friend- and stranger- associations because they are more difficult associations to form.

The contrasting roles of the self-related network and the attentional control network are further evidenced by neuropsychological studies of patients who have acquired damage to those brain regions. Sui, Enock et al. (2015) assessed the perceptual-matching performance and episodic memory of patient SC, who had damage to the left ventro-medial prefrontal (vmPFC) cortex and the insula, and patient RR, who had damage to the temporo-parietal (TP) cortex. Across the two tasks, SC demonstrated a hypo self-bias; that is, the self-association advantage demonstrated by neurotypical participants was eliminated. In contrast, RR demonstrated a hyper self-bias during perceptual matching, relative to healthy controls matched for age and education. However, RR's episodic memory reflected a normal self-bias effect.

To account for these results, Sui et al. (2015) suggested that SC's acquired lesion to the vmPFC disrupted access to their core self-representation. This argument coincides with other neuroimaging studies that found that the vmPFC is recruited when making self-related judgements (Kelley et al., 2002; Mitchell, Banaji, & Macrae, 2005; Sui, Rotshtein, et al., 2013). In contrast, the posterior superior temporal sulcus, a region damaged in RR, has been associated with self-biases in perceptual matching (Sui,

Rotshtein, et al., 2013). Sui et al., (2013) proposed that the pSTS mediates top-down attentional responses to socially relevant stimuli. Therefore, damage to the pSTS reduces the ability to allocate sufficient attentional resources when making the more difficult shape-label associations (i.e., to the stranger). However, with the reduction of attentional control, strong drivers of attention such as self-relevance go un-modulated, resulting in a hyper self-bias. As there are fewer attentional demands for the episodic memory task this hyper self-bias was not found.

Humphreys and Sui (2016) proposed a framework to explain the interaction between neural circuitry involved in self-processing and attentional control which they termed the Self-Attention Network (SAN). The model has three main nodes: (i) a general-purpose top-down attention control network, which recruits both the dorsolateral prefrontal cortex (DLPFC) and the intra-parietal sulcus (IPS), (ii) the self-representation, which is situated at the ventro-medial prefrontal cortex (vmPFC), and (iii) a node for the bottom-up orienting of attention which is housed at the posterior superior temporal sulcus (pSTS). Interactions between these three nodes determines performance. The vmPFC, thought to house self-representation, is rapidly triggered when self-relevant information is detected. Activation of the vmPFC in turn activates the pSTS node, priming it to respond to self-relevant information, resulting in self-biases in perception and attention. In addition, the top-down attentional control network operates to moderate the bottom-up activation of the vmPFC and the pSTS. For example, on non-matching trials in a perceptual matching task (i.e., if the self-associated shape is paired with the friend label) the fast triggering of a response to a self-associated stimulus could result in the participant responding incorrectly. To militate against this, the attentional control network deploys top-down inhibition of self-representations.

Much of the empirical support for interactions between the nodes come from fMRI data recorded while participants performed a perceptual matching task (Sui, Rotshtein, et al., 2013). Although there was strong correlational evidence to suggest an excitatory association between the VmPFC and the LpSTS, and an inhibitory association between the DLPFC and the vmPFC, evidence of a connection between the

attentional control network and the pSTS was limited. Indeed, Sui, Rotshtein, et al. (2013) found no evidence of a correlation between activation in the DLPFC and the pSTS under perceptual matching conditions. Therefore, further studies are needed to establish if a connection between the attentional control network and the pSTS modulates bottom-up attention to the self.

Summary: Neuroimaging studies of healthy individuals and individuals with acquired brain injuries have indicated three interacting neural 'nodes' that are involved in responding to self-relevant information. These nodes include (i) a top-down attention control network (DLPF/IPS), (ii) the house of self-representation (vmPFC), and (iii) a node for the bottom-up orienting of attention (pSTS). Together, these interconnected node make up the Self Attention Network (SAN, Humphreys & Sui, 2016).

1.9 Defining 'the self' in self-bias

So far it has been assumed that self-biases are universal in human cognition. However, research has found that both individual differences and cultural influences play an important role in the manifestations of self-bias. For example, a recent study found that individual differences in self-consciousness modulated the self-reference effect in memory (de Caso, Poerio, Jefferies, & Smallwood, 2017). Participants demonstrating a trait-like tendency for private self-consciousness (i.e., the tendency for introspection and the evaluation of internal beliefs and values) showed a stronger memory bias for self-relevant information. This finding was associated with individual differences in the functional organisation of the brain at rest. Furthermore, Humphreys and Sui (2015a) found that these individual differences in the magnitude of the self-bias were relatively stable over time.

One of the key factors thought to generate individual differences in self-bias is the culture that an individual has experienced. Most notably, people from Western cultures have typically been found to demonstrate stronger self-biases than people from

East Asian cultures (Liu et al., 2015; Sparks, Cunningham, & Kritikos, 2016; Sui, Hong, Hong Liu, Humphreys, & Han, 2013; Sui, Liu, & Han, 2009). This is thought to reflect differences in self-construals between Western and East Asian cultures (Markus & Kitayama, 1991). In the West there is a greater emphasis on self-focused attention and independence, whereas in East Asia 'the self' is more defined by social connections with other people (Gudykunst et al., 1996; Markus & Kitayama, 1991; Triandis, Bontempo, Villareal, Asai, & Lucca, 1988).

Sui, Liu and Han (2009) investigated whether cultural differences in self-processes influenced the recognition of self-faces relative to other-faces. British and Chinese students were asked to judge the orientation of their own face and a familiar face while the researchers recorded event-related potentials. The British students were faster to respond to own-face stimuli than other-face stimuli, and the own-face stimuli elicited a larger negative activity at 280–340ms over the frontal-central area (N2) relative to the familiar face. In comparison, the Chinese students' behavioural responses indicated a weaker self-bias, and this was reflected in a reduced anterior N2 amplitude to the own face compared with the familiar face. The results of this study suggest that culture modulates the neurocognitive processes of self-recognition that underlie the self-bias in face processing.

There is evidence that people show similar biases with information relating to groups they are members of as they do to information directly relating to the self. For example, Moradi, Sui, Hewstone and Humphreys (2015) asked participants to perform a perceptual matching task (Sui et al., 2012) in which they matched novel shape stimuli with the badge of their favourite football team (in-group), a rival team (out-group), and neutral teams. Participants were more efficient at responding to the shape stimuli associated with the in-group compared to an out-group stimulus (both rival and neutral outgroups). It is possible that people's self-representation extends to include all information associated with the self, including the groups that they belong to.

A recent study conducted by Enock, Sui, Hewstone and Humphreys (2018) directly investigated the relationship between self-prioritisation and in-group-

prioritisation in a sample individuals who are members of rowing teams. Participants performed two perceptual matching tasks. In one they associated geometric shapes with self, friend, and stranger labels and in the other they associated geometric shapes with labels referring to their own rowing team, a rowing team that they considered to be a rival, and another rowing team that they considered to be neutral to theirs, a non-rival. Participants were required to respond whether a presented shape stimulus had been paired with the appropriate label. Across the two tasks there was a reliable benefit of self-association and own-team-association. The data revealed a correlation between the size of the self and the own-team advantages at the level of the individual. Participants who reported a stronger sense of group identity demonstrated greater overlap in the perception of self- and own-team- associated shapes. This pattern of results is consistent with there being overlap between a participant's self-representation and representation of the group in which they belong. The data is also suggestive of a common mechanism driving self-biases and in-group biases in perceptual processes.

However, it is possible that people demonstrate more efficient responding to in-group stimuli because in-group stimuli are more likely to activate knowledge and affect than stimuli associated with a group that the participant is unaffiliated with. To disentangle these factors Allan, Morson, Dixon, Martin and Cunningham (2017) introduced participants to two virtual strangers; one who was designed to hold similar opinions to the participant, and one who was designed to hold opposing opinions to the participant. Allan et al. (2017) hypothesised that participants would have access to relatively more high-quality knowledge about the similar other than the dissimilar other, by activating their own self-representation. Participants then were presented with pairs of objects and were instructed to select one object either for themselves, or for the similar or dissimilar other. Following the selection task participants were given a surprise object recognition test, which also required them to identify for whom the recognised object had been chosen for. Recognition memory and source memory was superior for the self-owned items. Recognition was also better for the similar-other items than the dissimilar-other items. However, there was no difference in the accuracy of source memory for items assigned to the similar and dissimilar others. The results of this study

suggest that while the simulation of a similar other benefitted from some level of self-processing, there was still a processing distinction between stimuli associated with the self and stimuli associated with a similar other. Therefore, when it comes to cognition, being 'self-like' is not the same as being the self.

Taking the merging of self and other one step further, Payne, Tsakiris and Maister (2017) adapted Sui et al's (2012) perceptual matching paradigm so that instead of associating the self and others with shape stimuli, participants associated the self and other with images of gender-matched faces belonging to strangers. Response latencies suggested that participants were faster to associate an unfamiliar face with the self than to a friend or stranger. That is, a typical self-association bias was found. The authors were also interested in whether incorporating someone else's face within a self-representation alters one's physical self-representation. Before and after completing the perceptual matching task, participants were asked to rate face images for similarity to their own (real) face. Each image was constructed by morphing an image of the participant's own (real) face with that of one of the three images that had previously been associated with either themselves, the friend, or the stranger. No difference was found in similarity ratings across the morphed images for both the pre- and post-test. Therefore, following the successful association of an unfamiliar face and the self (as demonstrated using the perceptual matching paradigm) participants did not perceive their own face to become more similar with the assigned unfamiliar face. Thus, while representations of self and other can become associated, evidence suggests that they remain distinct

Summary: Do cognitive self-biases reflect a distinction between self and others? The answer to this question may depend on the cultural affiliation of the individual. For example, people from individualistic cultures show a more marked self-prioritisation effect than people from more collectivist (i.e., East Asian) cultures. Furthermore, evidence suggests that self-biases extend to stimuli associated with other types of 'in-groups', such as a sports team. However, it is not clear how affiliating the self with other individuals influences self-biases. Recent research has found associating the self with the

face of another individual alters the representation of the other individual, rather than the self.

1.10 Chapter Overview

This chapter provided an overview of research demonstrating self-biases across the cognitive domains of memory, attention, and perception. For example, people are better at recognising and recalling information encoded in reference to the self than in reference to other people (Klein, 2012) and these self-relevant memories typically incorporate greater conceptual and contextual detail (Sui, 2016). Furthermore, people show processing advantages for their own name compared to the names of others, and for their own face compared to the faces of others, indicating that self-relatedness also influences processes involved in the visual processing of social stimuli (Humphreys & Sui, 2016).

There is also a large body of evidence to suggest that self-relevance enhances processes of attention. However, there is contention as to whether the bias reflects the initial capturing of attention, or manifests later in the processing stream, such as the top-down modulation of attention to high-priority stimuli (Tacikowski & Nowicka, 2010). Humphreys and Sui (2016) proposed the Self-Attention Network (SAN) to account for the fact that research has indicated both that self-biases modulate top-down attentional processing *and* that self-biases can also be influenced by attention. The SAN describes the interaction between a top-down attentional control network (which includes the dorso-lateral prefrontal cortex (DLPFC) and the intra-parietal sulcus (IPS)) and a social salience network which responds to self-relevant information in a bottom-up fashion (involving the posterior superior temporal sulcus (pSTS) and the ventromedial prefrontal cortex (vmPFC)).

Although over fifty years of research has indicated that the self has a powerful and unique influence on cognition, it is possible that the purported self-bias is simply a

manifestation of a more generic cognitive bias, such as those relating to the inherent familiarity, reward value, or positive emotional valence of stimuli associated with the self. However, 'new wave' methodologies, such as Sui et al's (2012) perceptual matching paradigm, have allowed researchers to disentangle some of these factors. Most notably, people are found to be much more efficient at associating novel stimuli with the self than with other individuals. This robust finding suggests that self-biases are not limited to over-learned stimuli such as one's own face and name. Instead, novel stimuli are rapidly incorporated within the self-representation, resulting in the modulation of relatively low-level processes of perception and attention.

Furthermore, studies adopting the perceptual matching paradigm have found that although the self-association benefit appears to be qualitatively similar to that of high-reward or positive-emotion, these effects can be disassociated from one another, indicating that they are not fully equivalent (Stolte et al., 2017; Sui & Humphreys, 2015a). This leads us to believe that self-relevant stimuli might undergo unique processing compared to other high-salience stimuli. Indeed, neuroimaging studies have indicated both overlapping and distinct neural networks involved in the processing of self-relevance, reward, and positive emotion (de Greck et al., 2008; Enzi et al., 2009; Sui & Humphreys, 2015d). Thus, there is both behavioural and neuroimaging data in support of the idea that self-relevant information undergoes some level of distinct processing from other forms of salient information.

Finally, this chapter ended with a discussion of how self-biases reflect the demarcation of self from others. The smaller self-bias demonstrated by East Asians compared to Westerners (Liu et al., 2015; Sui & Humphreys, 2013b) suggests that culture plays an important role in how we define ourselves in relation to others. For instance, people from collectivist cultures may have greater overlap between their representations of self and other than people from cultures that promote a more individualistic self-construct (Markus & Kitayama, 1991). Furthermore, people have been found to have a similar bias for stimuli associated with an in-group as they do for stimuli associated with themselves alone (Enock et al., 2018; Moradi et al., 2015).

However, other research has indicated that self-representations are well protected and are not easily confused with representations of similar others (Payne et al., 2017).

2 SELF-OTHER CONTROL

Chapter 1 provided an overview of research demonstrating how information relating to the self is prioritised for cognitive processing over information relating to other individuals. Such research has been highly enlightening as to how the brain selects, processes, and organises incoming information – uncovering the significant role of the self in defining how we experience the world. However, the vast majority of experiments investigating cognitive self-biases have been conducted with individual participants using computer-based tasks. What is yet to be addressed is how information relating to the self and others are processed within the context of social interaction.

Humans are highly social creatures and, outside of the lab, many tasks are conducted in the presence of other people. Therefore, individuals are frequently required to coordinate their behaviours with those around them. Such interpersonal coordination is facilitated by the ability to 'mentalize', or theorise about, the internal mental states of the self and others (Frith & Frith, 2006b). In particular, social interaction often requires co-representation of goals, behavioural intentions, and action plans, between two or more co-actors (Sebanz, Bekkering, & Knoblich, 2006a; Vesper

et al., 2017). When co-representing a shared task with another person it is important that all task components are coded in a functionally similar way, so that the resulting behaviours correspond with one another (Prinz, 1990). It is also important that the task components associated with the self and other are distinguishable to avoid confusions of agency (i.e., knowing who is to execute what action).

The practice of sharing task representations with another individual is also challenged by the fact that human cognition is fundamentally ego-centric (See Chapter 1). If self-relevant information is always prioritised over other-relevant information, it would be difficult (if not impossible) to take the perspective of someone else. However, it is well-established that most individuals are able to cognitively 'put themselves in someone else's shoes' (Apperly, 2010; Surtees & Apperly, 2012). It therefore follows that there exists a top-down mechanism that can carefully balance activation of self- and other- representations during social interaction (de Guzman, Bird, Banissy, & Catmur, 2016; Sowden & Shah, 2014). If the ego-centric bias is left unchecked, then individuals will ascribe their own beliefs and intentions on to others. However, if self-prioritisation is completely supressed then individuals will be unable to separate their own action plans from those of their co-actor. Recent research has begun to investigate and identify the neural mechanisms of self-other control when performing tasks with other individuals (de Guzman et al., 2016; Sowden & Shah, 2014). Henceforth the term self-other control will refer to the control of neural representations of the self and of other people (see also Sowden & Shah, 2014).

The aim of Chapter 2 is to review evidence of self-other control processes across socio-cognitive functions such as imitation, perspective taking, theory of mind, joint action, and task sharing. Furthermore, I will look to neuropsychological studies of individuals with disorders reflecting impairments of self-other control; such as autism spectrum disorder, schizophrenia, and mirror-touch synaesthesia. Such studies have been used to identify the neural substrates of mentalising, which remains so poorly understood that it has been dubbed the 'dark matter' of social neuroscience (Przyrembel, Smallwood, Pauen, & Singer, 2012).

2.1 Imitation

There is a large body of research indicating that humans have an automatic tendency to imitate the actions of others (Heyes, 2011). This tendency is demonstrated from infancy, with new born babies imitating the facial expressions of their caregivers (Meltzoff, 1990; Vincini, Jhang, Buder, & Gallagher, 2017). The proclivity for imitation continues in to adulthood. Evidence for such imitation typically comes from stimulus-response compatibility tasks in which participants observe another individual performing an action which is either compatible or incompatible to the one required by their own task (Brass, Bekkering, Wohlschläger, & Prinz, 2000; R. Cook, Bird, Lünser, Huck, & Heyes, 2011; de Guzman et al., 2016; Meltzoff & Decety, 2003; Santiesteban, White, et al., 2012). For instance, Stürmer et al. (2000), presented participants with a colour cue, directing them to either open or close their hands. This colour cue was superimposed over a task irrelevant video of another individual performing one of the two hand gestures. Participants were quicker to initiate the correct action when the observed action was response-compatible than when it was response-incompatible.

How and why does imitation occur? Understanding how imitation is possible requires solving the 'correspondence problem' (Brass & Heyes, 2005). That is, how are people able to convert a representation of an observed action to a motor representation of the same action that can then be executed? Current theory proposes that the answer to this question resides in the 'mirror neuron system' (MNS), which is comprised of a number of cortical regions that are activated both when an action is observed, and when the action is executed (Gallese & Goldman, 1998; Iacoboni, 2009; Oztop, Kawato, & Arbib, 2006). This correspondence is thought to indicate that observed and performed actions are coded in a functionally similar manner (Prinz, 1990). Accordingly, simply observing an action will then co-activate the corresponding motor representation, resulting in the imitative response (Borroni & Baldissera, 2008; Fadiga, Fogassi, Pavesi,

& Rizzolatti, 1995; Gangitano, Mottaghy, & Pascual-Leone, 2001; Heyes, 2011; Oztop et al., 2006).

Support for this theory of automatic imitation comes from studies involving single-pulse transcranial magnetic stimulation (TMS), a non-invasive procedure that uses magnetic fields to stimulate neurons in the brain. When TMS is used to disrupt activation in the ventral premotor cortex (a putative MNS region) people no longer demonstrate 'motor resonance', the subthreshold activation of an imitative response when observing an action (Avenanti, Bueti, Galati, & Aglioti, 2005). Furthermore, studies have found that TMS disruption of the same region actually reduces automatic imitation (Catmur, Walsh, & Heyes, 2009; Mengotti, Ticini, Waszak, Schütz-Bosbach, & Rumiati, 2013).

While there is much evidence to suggest that observing an action automatically activates corresponding motor activation in the observer, it is not the case that people get stuck in a perpetual loop of imitation. Therefore, individuals must be able to distinguish between representations of an observed action and their own intended actions, in order to prioritise the latter over the former. How then, do representations of an observed and an intended action become uncoupled? And how are individuals able to prioritise their own goal-driven behaviour over the tendency to imitate? The answers to these questions likely involve some form of self-other control mechanism (Sowden & Shah, 2014).

Neuropsychology studies have found that patients with brain lesions to the frontal lobe often demonstrate excessive imitative behaviours and reduced imitative control (Spengler, von Cramon, & Brass, 2010). This finding suggests that there is a control mechanism residing within (or at least recruiting) the frontal lobe, which facilitates the inhibition of unwarranted imitation. Interestingly, evidence suggests that the control of imitation does not rely on the same mechanisms involved in other tasks that require more general inhibitory control, such as the Stroop task (Brass, Derrfuss, Matthes-von Cramon, & von Cramon, 2003; Brass & Heyes, 2005) or spatial compatibility tasks (Cross, Torrisi, Reynolds Losin, & Iacoboni, 2013). This dissociation

indicates that there is a unique control mechanism for imitative control, possibly reflecting the social processes involved in imitation, including the involvement of the MNS (Cross & Iacoboni, 2014).

In support of this theory, there is evidence that the control of imitation inhibition is related to social-cognitive processes involved in distinguishing between the self and others (Brass, Ruby, & Spengler, 2009; Spengler, von Cramon, et al., 2010). For example, in a study of patients with frontal lobe lesions, a positive correlation was found between the ability to control imitative tendencies and the ability to attribute mental states (i.e., knowledge, beliefs, intents, and emotions) to the self and to others (Spengler, von Cramon, et al., 2010). The same study found that patients with lesions to the temporo-parietal junction (TPJ) demonstrated a positive correlation between imitative control and perspective taking abilities (both visual and conceptual). These correlations remained significant even when executive function was controlled for, indicating the functional specificity of the relationship. Therefore, the results of this study suggest that the underlying processes involved in imitative control have more to do with the control of representation of self and other, than more general inhibitory control.

Importantly, the link between inhibitory control and processes of self-other distinction are not merely correlational. Santiesteban et al. (1988) trained different groups of participants using tasks that promoted either: (i) the imitation of an observed action, (ii) the inhibition of imitation of an observed action, or (iii) general (non-social) inhibitory control using a stroop-like task. The group who were trained with the inhibition of imitation task showed improved performance on a visual perspective taking task. However, there was no such improvement for the other two groups. This pattern of results indicates that the same processes of self-other control underlie imitation inhibition and perspective taking. To further investigate the role of self-other control in social cognition, the next section will consider the processes involved in perspective taking in greater detail.

Summary: Humans demonstrate an automatic tendency to imitate others. The ability to translate an observed action into a motor representation that can be executed

is thought to be due to the 'mirror-neuron' system. However, it is not the case that neurotypical individuals imitate every action they observe. Neuroimaging and behavioural data has indicated that there are specific self-other control mechanisms involved in the inhibition of imitative tendencies.

2.2 Perspective Taking and Theory of Mind

Theory of mind (TOM) is an umbrella term for the ability to understand how another person's mental states (such as their beliefs, knowledge, and intentions) may differ from one's own (Baron-Cohen, Leslie, & Frith, 1985; C. D. Frith & Frith, 2005; Perner, 1999). The term most broadly denotes the ability to take another person's perspective, both visually and conceptually. Being able to relate one's own thoughts to those of another person is a fundamental component of most intersubjective transactions.

Unlike the imitative behaviours that humans demonstrate from birth, TOM abilities do not typically emerge until the age of four years old (Wellman, Cross, & Watson, 2001). The delay in acquisition relative to imitation likely reflects the relative complexity of the processing involved in TOM. Unlike imitation, TOM requires deliberate 'mentalising' about the thoughts, intentions, or beliefs of another individual (Gallese & Goldman, 1998; Meltzoff & Decety, 2003; Santiesteban, White, et al., 2012). TOM abilities have traditionally been examined using false belief tasks. For example, Gopnik and Astington (1988) presented young children with a box of chocolates that had recognisable packaging. When the children opened the box they found, to their surprise, that it contained pencils rather than chocolates. The children were then asked what they thought a friend would believe was inside the box before they could open it. Most of the children aged three-years-old or younger reported that the friend would know that the box contained pencils. In contrast, most children aged four years and above realised that their friend would hold the same false belief that they once did (i.e.,

that the box contained chocolates). It has been proposed that children younger than three years old have not yet developed the capacity to understand that the beliefs of others can differ both from reality and from their own situational knowledge (U. Frith & Frith, 2003). Alternatively, it might be that the younger children were able to take the perspective of their friend, but they were unable to prevent their own knowledge about the situation from dominating (Carlson & Moses, 2001; Müller, Zelazo, & Imrisek, 2005; Surtees & Apperly, 2012).

One of the key challenges in TOM and perspective taking is to prevent self-representations from overriding representations of the mental states of others (Epley, Keysar, Van Boven, & Gilovich, 2004; Epley et al., 2004; Galinsky & Mussweiler, 2001; Surtees & Apperly, 2012). There is now converging evidence to suggest that executive function plays a crucial role in successful visual perspective taking. For example, children's performance on Stroop-like tasks, which require a high degree of cognitive control, has been found to be positively correlated with their perspective taking abilities (Carlson, Moses, & Breton, 2002; Fizke, Barthel, Peters, & Rakoczy, 2014; Rakoczy, 2010). Also, TOM abilities such as perspective taking and emotion recognition, have been found to positively correlate with executive functions relating to inhibitory control and working memory in both young (18-30 years) and older (60+ years) adults (Bradford, Brunsdon, & Ferguson, 2017).

Neuropsychological studies of patients with specific brain lesions have been valuable in investigating the relationship between executive functions and perspective taking. In particular, previous work has found that patients with acquired brain injuries to the right lateral prefrontal cortex, resulting in deficits in executive function, also demonstrated deficits in mentalising abilities involved in perspective taking and TOM (Aboulafia-Brakha, Christe, Martory, & Annoni, 2011; Apperly, Samson, Chiavarino, Bickerton, & Humphreys, 2007; Qureshi, Apperly, & Samson, 2010a; Samson, Apperly, Kathirgamanathan, & Humphreys, 2005). Samson, Housthuys and Humphreys (2015) investigated whether the domain-specificity of the executive function deficit is important. Pairs of brain damaged patients, who had similarly located lesions to the prefrontal cortex, played a card game against each other. In the game each participant

drew a coloured card from a deck (either red or blue). Participants were told that a third card would then be drawn from the deck. On 'win' trials participants would receive a reward if the third card was the same colour as their own card. On 'lose' trials participants would incur a loss if the colour was the same as their card. Prior to the third card being revealed, participants were asked to report either: (i) the colour they desired the card to be, or (ii) the colour their opponent would desire the card to be.

This task required executive control on two fronts. First, in order to make a judgement about the opponent's desires, the patients must inhibit their own (salient) desires. Second, in order to make an avoidance decision (i.e., how to avoid a loss) patients must supress the tendency to make an approach decision (i.e., how to win a reward). The results of the experiment demonstrated a classic double dissociation between self-perspective inhibition and the ability to resist approach motivation. One pair of participants demonstrated impairments in supressing their own desires in order to report the desires of the opponent but showed little difficulty in selecting between an avoidance response and an approach response. The other pair of patients showed the reverse pattern: they were able to suppress their own desires when taking the perspective of others but were biased towards making an approach response even on loss trials, resulting in errors. The results of this study suggest that the specific impairment in inhibiting one's own perspective cannot be accounted for by a general deficit in inhibitory control.

A further study found evidence for a dissociation between self-perspective inhibition and more general executive functions, but within a neurotypical population (Qureshi et al., 2010a). Participants were required to perform a visual perspective taking task while also completing a secondary task that increased the demands on executive function (i.e., they had to tap their index finger twice whenever they heard a single tone being played, or once if they heard two tones being played). The results revealed that even under dual task conditions, participants were able to take the perspective of another person. However, when the two perspectives significantly differed, the participants had difficulty selecting the appropriate perspective when required. This

pattern of results indicates a distinction between perspective taking and self-other control processes.

Across studies, first-person perspective (i.e., a self perspective) has been found to interfere with representations of a third-person perspective, more so than the other way around. While this may be due to the inherent salience and prioritisation of self-relevant representations, a potential confound has been identified. It has been proposed that taking a third-person perspective is inherently more difficult than taking a first-person perspective because it requires 'mental rotation' (Mattan, Quinn, Apperly, Sui, & Rotshtein, 2015). To address this problem, Mattan et al. (2015), asked participants to take the perspective of avatars that either represented themselves (self-relevant third-person-perspective) or a best friend (other-relevant third-person-perspective). Therefore, the authors could directly investigate whether self-representations were prioritised over other-representations during perspective taking, without the confound of task difficulty. The avatars existed within a virtual room, and either had an unobstructed 'view', or partially obstructed 'view', of an array of dots presented on one wall. On each trial participants were asked to report how many dots a selected avatar could see. Overall, participants demonstrated an advantage for judgements about the self-avatar's perspective than the other-avatar's perspective. The results of Mattan et al's study suggest that self-relevant perspectives are prioritised over other-relevant perspectives even when in the third person. This speaks to the intrinsic salience of self-representations (See Chapter 1).

Summary: The ability to theorise about and represent the internal mental states of others is an essential part of successful social interaction. However, given that human cognition is intrinsically ego-centric, top-down control is needed to prevent the ascription of one's own knowledge and perspective on to other people. Behavioural and neuroimaging studies have indicated that the inhibition of self-representations is a distinct process from general executive control.

2.3 Empathy and Theory of Mind

Empathy describes the subjective understanding of another individual's emotions (Decety & Jackson, 2004). The socio-affective processes involved in empathy are thought to be largely distinct from the socio-cognitive processes in involved in theory of mind and perspective taking (Preckel, Kanske, & Singer, 2018). Recent neuroscientific research has indicated that processes of empathy are based on activity within different neural networks than those involved in TOM and perspective taking (Kanske, Böckler, Trautwein, & Singer, 2015). Nevertheless, both processes are jointly important when attempting to understand and interact with others. One of the key commonalities between processes of empathy and TOM is that they both require careful self-other control. Maladaptive self-other distinctions during empathy can result in severe emotional distress (Decety & Lamm, 2009; Eisenberg & Eggum, 2009), whereas successful self-other control allows individuals to experience compassion for one another (Singer & Klimecki, 2014).

The first studies to investigate the neural substrates of empathy focused on empathy for pain (Jackson, Rainville, & Decety, 2006; Singer et al., 2004). Both directly experiencing pain and observing the pain of others was found to activate the anterior insula (AI) and anterior middle cingulate cortex (aMCC). Subsequent meta-analyses have identified these two regions as part of a core network that is activated both when experiencing pain and when witnessing the suffering of others (Y. Fan, Duncan, de Greck, & Northoff, 2011; Lamm, Decety, & Singer, 2011). This common activation has been taken as support for simulation theories and theories of embodied social cognition, which propose that people activate their own emotion system to understand and simulate what another person is sensing or feeling (Decety & Jackson, 2004; Gallese, 2001). However, it must be noted that there is considerable controversy around the idea that shared neural activation is sufficient evidence for shared mental representations (Lamm & Majdandžić, 2015).

The neural correlates of empathy and TOM have typically been studied in isolation from one another. However, a recent study that implemented a novel task called the EmpaToM, which independently manipulated both TOM and empathy, found that individual difference in ToM and empathy were not correlated at both neural and behavioural levels (Kanske et al., 2015). Therefore, it is not necessarily the case that proficient mentalisers are particularly strong empathisers. Despite being separable, there are certain paradigms that co-activate processes of empathy and TOM. For instance, activation within the TOM network was found when participants were presented with a cue that *suggests* someone else is in pain, but not when participants were presented with a visual depiction of someone experiencing pain (Lamm et al., 2011). This pattern of results likely reflects the fact that both cue-based empathy paradigms and ToM paradigms require fairly abstract reasoning about the cognitive/affective states of others. In contrast, the pictorial depiction of someone experiencing pain activates the action-observation network that helps them to simulate (to a highly attenuated extent) what it would feel like to go through a similar experience.

Furthermore, there is evidence to suggest that during complex interactions, networks involved in empathy and TOM interact with one another (Preckel et al., 2018). In particular, when individuals are confronted with others' distress, activation within the AI, which reflects the 'sharing' of negative affect, inhibits activation within the TPJ, one of the key regions involved in mentalizing (Kanske, Böckler, Trautwein, Parianen Lesemann, & Singer, 2016). It has been proposed that the inhibition of activity in the TPJ when faced with a distressing situation is thought to prepare the individual for immediate action (Menon & Uddin, 2010). These stress-related mentalizing deficits have also been implicated as an important component of psychopathology, particularly in the context of borderline personality disorder (Fonagy & Luyten, 2009).

The ability to distinguish between self and other is just as crucial for empathy as it is for TOM. Failure to distinguish and control activation between one's own affective state and the affective state of others can lead to a melding of the two states. This can result in either an ego-centric bias, the projection of one's own emotions on to someone else (Royzman, Cassidy, & Baron, 2003; Silani, Lamm, Ruff, & Singer, 2013), or an

allocentric bias, whereby the emotional states of others overly influence one's own emotions (Hoffmann et al., 2016). The suppression of emotional egocentricity has been associated with enhanced activation within an anterior temporoparietal region in the right supramarginal gyrus (rSMG). When activity within the rSMG was disrupted using transcranial magnetic stimulation, emotional egocentricity was enhanced (Silani et al., 2013). Furthermore, increased coupling between the rSMG and the dorsolateral PFC has been associated with a reduced emotional egocentric bias (Steinbeis, 2016). Therefore, it appears that the rSMG and the dorsolateral PFC both play an important in managing the egocentric bias when empathising.

A relationship has been found between the ability to empathise with others and the likelihood of participating in pro-social behaviours (Tusche, Böckler, Kanske, Trautwein, & Singer, 2016). Impairments of self-other control can have serious negative behavioural consequences when social interactions require empathy. An egocentric bias has been associated with antisocial behaviour such as hyper-aggression (Winter, Spengler, Bermpohl, Singer, & Kanske, 2017). In contrast, an allocentric bias has the outcome of empathetic distress; the phenomenon that people experience personal suffering when they observe the suffering of others; which can in turn can have asocial or antisocial behavioural consequences (Klimecki, Vuilleumier, & Sander, 2016).

Summary: Empathy is the ability to understand the mental states of others. The socio-affective processes involved in empathy have distinct neural substrates from the socio-cognitive processes involved in TOM. However, processes of self-other control are equally as important for empathy as they are for TOM. Impairments of self-other control can manifest as an ego-centric bias (resulting in the projection of one's own affective states on to others) or an allocentric bias (being overly influenced by the affective states of others). Careful self-other control during empathy plays an important role in promoting pro-social behaviour.

2.4 Relationship Between Executive Functions and Self-Other Control

Executive functions are a set of high-level cognitive processes, associated with the frontal lobes, that control lower level processes in the service of goal-directed behaviour (Friedman & Miyake, 2017). Executive functions include the inhibition of automatic responses (Brass et al., 2009), resisting distraction (Krabbe, Ellbin, Nilsson, Jonsdottir, & Samuelsson, 2017), task switching (Davidson, Amso, Anderson, & Diamond, 2006), working memory (Kirova, Bays, & Lagalwar, 2015), and cognitive flexibility (see Diamond, 2013 for a review of executive functions).

The ability to simultaneously represent and manipulate between potentially conflicting mental representations of the self and other (as required during empathy, theory of mind, and perspective-taking) is thought to employ executive functions (Decety & Sommerville, 2003a). In fact, studies have found that the development of executive function during childhood is associated with the acquisition of theory of mind (Aboulafia-Brakha et al., 2011; Carlson, Claxton, & Moses, 2015; Carlson et al., 2002; Hala, Hug, & Henderson, 2003). There is also robust evidence to link theory of mind with specific executive function subcomponents; such as inhibition (Carlson, Moses, & Claxton, 2004), attention shifting (Leslie, Friedman, & German, 2004), and working memory updating (Austin, Groppe, & Elsner, 2014). This makes intuitive sense, as attributing mental states to others requires (1) the temporary inhibition of one's own self representation, (2) shifting attention from representations of oneself to representations of the other individual, and (3) updating the representation of the other individual. There is also evidence to suggest that poor executive functioning is associated with decreased empathy and increased antisocial behaviours during childhood (Decety & Jackson, 2004), suggesting that executive functions may play an important role in developing an understanding and other individual's emotional and mental states.

A study by Qureshi, Apperly and Samson (2010b) on visual perspective taking found that executive functioning played a role in the selection between two differing

perspectives (self and other), but not the ability to represent another individual's perspective. In this study participants were asked to perform a level-1 visual perspective taking task (i.e., they reported how many items present in a room were visible to the self and/or a positioned avatar) and a concurrent secondary task which increased demands on executive function (tapping a finger twice if one tone was played, or once if two concurrent auditory tones were played). Reporting of the avatar's perspective was disproportionately affected by the increased demands on executive function when the avatar held a *different* visual perspective from the self, but not when the avatar held the *same* visual perspective as the self. This finding suggests that attributing mental states to others operates independently of executive function, but that executive function is deployed when supressing one representation in favour of another.

Neuroimaging studies have shown that the medial prefrontal cortex (mPFC), a frontal area of the brain associated with executive function, is also activated during tasks that involve judging agency, perspective-taking, theory of mind, and empathy (Decety & Lamm, 2007; Decety & Sommerville, 2003a; Sperduti, Delaveau, Fossati, & Nadel, 2011). It has been proposed that the mPFC is activated during tasks that require the manipulation of self and other representations (i.e.,self-other control) because it facilitates the differentiation of task-relevant from task-irrelevant representations (Nicolle et al., 2012; Cook, 2014).

Summary: Executive Functions are cognitive processes involved in behavioural and cognitive control. The ability to successfully distinguish between and manipulate representations of the self and of others (as required in empathy, theory of mind, and perspective taking) relies on executive functions (such as inhibitory control, working memory, and selective attention), recruiting areas of the frontal lobe associated with cognitive control. However, the ability to form representations of the self and other does not rely on executive functions. Thus, executive functions are required for self-other *control* but no self-other *representation*.

2.5 Disorders of Self-other control

A number of clinical and sub-clinical conditions have been associated with the disruption of self-other control. The disorder that has received the most attention from researchers is autism spectrum disorder (ASD). Although ASD is considered to be a highly heterogenous disorder, the most recent revision of the Diagnostic Statistical Manual has grouped all symptomology into two high-order categories: (i) difficulties in social communication and interaction, and (ii) restrictive or repetitive behaviours (American Psychiatric Association, 2013). It is particularly striking that ASD has been associated with impairments of all socio-cognitive and socio-affective abilities that have so far been discussed in Chapter 2; including deficits in imitation inhibition (Spengler, Bird, & Brass, 2010; Vanvuchelen, Roeyers, & De Weerdt, 2011), perspective taking (Rehfeldt, Dillen, Ziomek, & Kowalchuk, 2007; Zwickel, White, Coniston, Senju, & Frith, 2010), TOM (Baron-Cohen, 1997; Williams, 2010) and empathy (Baron-Cohen & Wheelwright, 2004; Lombardo, Barnes, Wheelwright, & Baron-Cohen, 2007).

Due to individuals with ASD showing widespread difficulties in relating to others, it has been proposed that ASD can be explained as a dysfunction of the mirror-neuron system (MNS). This is known as the 'Broken Mirror Theory' of autism (Dapretto et al., 2006; Iacoboni & Dapretto, 2006; Ramachandran & Oberman, 2006). However, this theory is highly controversial and there is much evidence to suggest that individuals with ASD do, in fact, have an intact mirror-neuron system (Y.-T. Fan, Decety, Yang, Liu, & Cheng, 2010; Schulte-Rüther et al., 2015; Schulte-Rüther et al., 2017; Southgate & Hamilton, 2008; Sowden, Koehne, Catmur, Dziobek, & Bird, 2016). For instance, it is not the case that people with ASD are unable to imitate others, as would be expected with a dysfunctional MNS. In fact, automatic imitative behaviour found to be enhanced in autism relative to neurotypical controls (Sowden et al., 2016).

A more promising conceptualisation of the social dysfunctions associated with autism is an impairment of the top-down control or modulation of social behaviour (Castelli, Frith, Happé, & Frith, 2002; J. Cook, Barbalat, & Blakemore, 2012; Neumann,

Spezio, Piven, & Adolphs, 2006; Sowden & Shah, 2014). For instance, it has been found that in neurotypical populations, pro-social priming can have a modulatory effect on the automatic tendency to imitate others, but this effect is not found for individuals with ASD (J. Cook & Bird, 2012). Furthermore, deficits in imitative inhibition has been found to positively correlate with the severity of ToM impairments found in individuals with ASD (Spengler, Bird, et al., 2010).

Taken together, these studies suggest that individuals with autism have a disruption within the neural network involved in the top-down control of representations of 'Self' and 'Other' (Sowden & Shah, 2014). This could account for both hyper-imitative tendencies and difficulties in mentalising about the self and others. Lombardo and colleagues (2011) found evidence that this impairment resides in the right TPJ, an area that plays a crucial role in social cognition. A more recent study compared resting-state fMRI data from autistic and neurotypical adolescents and found no evidence that the organization of the TPJ significantly altered in ASD (Igelström, Webb, & Graziano, 2017). However, they did find disturbed cerebellar input to the right dorsal TPJ for individuals with ASD. Overall, neuroimaging evidence to date suggests that the TPJ is crucial component of a neurocognitive mechanism that is impaired in ASD.

Another, much rarer, disorder associated with impairments of self-other control is mirror-touch synaesthesia. Individuals with this disorder experience overt somatic sensations when they observe the tactile experiences of another individual (Banissy & Ward, 2013; Blakemore, Bristow, Bird, Frith, & Ward, 2005). The disorder is also associated with heightened empathy and emotion perception (Banissy & Ward, 2007; Fitzgibbon, Giummarra, Georgiou-Karistianis, Enticott, & Bradshaw, 2010). A proposed explanation for mirror-touch synaesthesia is the overactivity in the neural network that responds to the observation of touch.

In an fMRI study, Blakemore et al. (2005) compared brain activity of a mirror-touch synesthete with that of healthy controls, while they either observed another person receiving touch, or they received touch themselves. For both the synesthete and

the controls, common brain regions were recruited when observing and experiencing touch (such as the primary and secondary somatosensory cortex, premotor regions, and the superior temporal sulcus). However, relative to the controls, the synesthete showed hyperactivity when observing touch. Of note, mirror-touch syneasthetes only experienced a strong tactile sensation when they observed touch to a real body. When they observed touch to dummy bodies, pictures of bodies, or disconnected dummy body parts, their tactile sensations were significantly weaker (Holle, Banissy, Wright, Bowling, & Ward, 2011). The results of this particular study suggest that mirror-touch synaesthesia is not entirely bottom-up driven. Instead, knowledge about the experience of the recipient modulates the intensity of the response in a top-down manner.

Banissy, Kadosh, Maus, Walsh and Ward (2009) proposed a neurocognitive model of mirror-touch synaesthesia which consisted of three interacting mechanisms: (i) a 'what' mechanism that can discriminate between a real human and an object, (ii) a 'who' mechanism that discriminates between the tactile experiences of the self and other, and (iii) a 'where' mechanism that determines what part of the body is experiencing touch. They suggested that synaesthetes over-incorporate information about other's tactile experiences into their own body representations - a specific impairment of the 'who' mechanism. Therefore, mirror-touch synaesthesia is also considered to be a disorder of self-other control.

It has also been proposed that much of the symptom profile of schizophrenia can be explained by an impairment of self-other control (Sowden & Shah, 2014). For instance, individuals with schizophrenia often hold beliefs that thoughts have been implanted in their brain, or that events caused by someone else were caused by themselves (Prikken, van der Weiden, Kahn, Aarts, & van Haren, 2018). Delusions and hallucinations may reflect an inability to recognise their own thoughts as being internally generated, and therefore attribute the thoughts to an external source (Farrer et al., 2004). Studies have also found that patients often 'sub vocally' produce the voices they claim comes from an external source (Cho & Wu, 2014; C. D. Frith, 2014). Individuals with schizophrenia are also often easily distressed when observing the distress of others, suggesting that they feel the distress as their own (Lehmann et al., 2014; Montag, Heinz,

Kunz, & Gallinat, 2007). Furthermore, performance on TOM tasks (i.e., false belief tasks) has been found to be a good predictor of the severity of social dysfunction of people with schizophrenia (Abu-Akel & Shamay-Tsoory, 2013; Biedermann, Frajo-Apor, & Hofer, 2012; Bora & Pantelis, 2013). Therefore, the impaired social functioning associated with schizophrenia likely reflects a difficulty in distinguishing between the mental states of the self and the mental states of others.

Jardri et al., (2011) developed an fMRI procedure to examine the neural correlates of self-other distinction during speech exchange for individuals with schizophrenia and aged-matched healthy controls. The data revealed increased overlap between the self and non-self cortical maps in schizophrenia, relative to the controls. The authors proposed that increased overlap between self and non-self cortical maps is the neuro-physiological signature of schizophrenia. Therefore, both behavioural and neuroimaging data suggests that social dysfunction in schizophrenia results from a lack of distinction between self- and other- reflective processes.

Summary: Impairments of self-other control have been associated with a number of clinical and subclinical disorders of social functioning; namely ASD, mirror-touch synaesthesia, and schizophrenia. Neuroimaging and behavioural studies have indicated that individuals with these disorders can represent the cognitive states of others, but that they have difficulties in relating and distinguishing between representations of self and other. Such studies highlight the importance of sensitive self-other control for successful social functioning.

2.6 Neural basis of self-other control

So far, a number of studies have been discussed that use neuroimaging techniques to investigate the neural underpinnings of self-other control processes. This section will focus on two areas of the brain thought to be particularly involved in self-

other distinction: the temporo-parietal junction (TPJ) and the supramarginal gyrus (SMG) (Steinbeis, 2016).

Decety and Lamm (2007) conducted a large scale meta-analysis of over seventy neuroimaging studies that look at various aspects of social cognition. They found that the right inferior parietal cortex, at the junction with the posterior temporal cortex, plays a fundamental role in high-level socio-cognitive processes related to TOM and empathy. The authors also found that the TPJ is associated with lower-level (bottom-up) processes associated with having a sense of agency, and the re-orienting of attention to salient stimuli. They proposed that the most parsimonious interpretation of these findings is that complex socio-cognitive activities recruiting the TPJ also rely on lower-level computations that generate, test, and correct internal predictions relating to sensory events.

The TPJ region is large and has heterogenous projections; thus it has been parcelled into three distinct regions, which each have unique resting state functional connectivity profiles (Mars et al., 2012). It is the posterior TPJ that has the greatest connectivity with regions of the brain associated with the attribution of mental states; such as the medial pre-frontal cortex and the posterior cingulate (C. D. Frith & Frith, 2006b). Importantly, within the same set of participants, the posterior TPJ had increased activation both when inhibiting imitative behaviours and when attributing mental states to the self and others (Spengler, von Cramon, & Brass, 2009).

Causal evidence for the role of the TPJ in self-other control comes from studies using transcranial direct current stimulation (tDCS; Santiesteban, Banissy, Catmur, & Bird, 2012). Excitatory (anodal) stimulation of the TPJ improved both visual perspective taking (which involves prioritising other-representations over self-representations) and the ability to inhibit imitative tendencies (which involves prioritising self-representations over other-representations). However, excitatory stimulation of the TPJ did not change participants' ability to make social judgements that did not involve resolving conflict between representations of the self and other. This pattern of results suggests that the

TPJ has a specific and crucial role in processes of self-other control; such as reducing ego-centricity and supressing imitative tendencies.

A second region of the temporo-parietal area that is located more anteriorly than the TPJ, the supramarginal gyrus (SMG), has also been implicated in processes of self-other control (Silani et al., 2013). The authors found that the SMG was particularly active when participants were attempting to overcome their egocentricity bias when it came to understanding that another individual's emotional state differs from their own. The temporary disruption of the right SMG using repetitive transcranial magnetic stimulation was found to increase the egocentricity bias. Unlike the TPJ, which has a more general function in self-other control, the SMG appears to have a more specific role in relating and distinguishing the affective states of the self and other. This functional differentiation is supported by the finding that the SMG has the stronger connections to brain regions associated with empathetic responses (i.e. anterior insula, anterior cingulate cortex) than the TPJ (Mars et al., 2012). Therefore, different regions of the temporo-parietal areas have different roles for different social and cognitive domains requiring self-other control.

Summary: Self-other control has been associated with two key areas of the brain; the temporo-parietal junction (TPJ) and the supramarginal gyrus (SMG). The TPJ is found to play a crucial role in self-other distinction during perspective taking and processes of theory of mind. The SMG is thought to play a more specific role in relating and distinguishing the affective states of the self and other.

2.7 Joint Action and Agency

In daily life it is often beneficial to perform actions jointly with someone else; whether it be to divide labour or to share expertise. In social psychology joint action has been defined as "any form of social interaction whereby two or more individuals coordinate their actions in space and time to bring about a change in the

environment"(Sebanz, Bekkering, et al., 2006a, p. 70). Coordinating one's own actions with the actions of others to achieve a joint outcome, such as moving a heavy piece of furniture from one room to another, requires some kind of interlocking of behaviours, action plans, perceptions, and intentions (Knoblich, Butterfill, & Sebanz, 2011).

There are two possible forms of coordination involved in joint action; planned coordination, and emergent coordination (Sebanz, Bekkering, et al., 2006a). Planned coordination, at a minimal level, involves agents cognitively representing the desired outcome of the joint action, as well as the role they must play in achieving that outcome (Vesper, Butterfill, Knoblich, & Sebanz, 2010). There can be great variation in what is specified in representations of the other agent's task. For instance, an agent may attempt to predict the motives, thoughts and perspectives of their action partner; or an agent may simply expect an action partner to follow their lead (Knoblich et al., 2011).

In contrast to planned coordination, emergent coordination is based on direct action-perception couplings and can occur spontaneously without intention. For example, pedestrians have a tendency to synchronise their walking patterns (van Ulzen, Lamoth, Daffertshofer, Semin, & Beek, 2008) and people engaged in conversation tend to synchronise their body sway (Hari, Himberg, Nummenmaa, Hämäläinen, & Parkkonen, 2013). Such emergent coordination can occur in conjunction with, or independent from, joint plans or shared knowledge (R. C. Schmidt, Fitzpatrick, Caron, & Mergeche, 2011). Furthermore, two or more agents may independently respond to the same perceptual and motor cues in a similar manner, resulting in them unintentionally acting as a single coordinated entity (Marsh, Richardson, & Schmidt, 2009).

As has been previously discussed within the context of imitative behaviour, there is evidence to suggest that one's own actions and the actions of others are co-represented using the same neural resources (Knoblich et al., 2011; Rizzolatti & Sinigaglia, 2010; Vesper et al., 2017). For instance, van Schie, Mars, Coles and Bekkering (2004) demonstrated that when individuals observe someone else committing an action error, their medial frontal cortex shows similar activation to when the observer is

committing the error themselves. In an extension of the ideomotor principle (James, 1890), the 'common coding' theory of action claims that all actions, be they self or other generated, are represented by their perceivable events (Prinz, 1990). Accordingly, a common action representation can be shared between the executer of the action and any observers.

The ability for multiple individuals to hold a shared action representation is considered to be a cornerstone of social interaction (Sebanz, Bekkering, et al., 2006a). However, the idea that an individual represents another person's actions in a functionally equivalent way to their own raises important questions as to the attribution of agency. How is it possible to distinguish one's own action from the action of another individual if they are represented in the same way? Take the example of two individuals performing a piano duet. In order to synchronize the timing of their playing, the pianists must monitor the actions of their partner while integrating their own actions. How would this be possible if representations of self- and other- generated actions were functionally equivalent? The equivalence between the two representations would cause ambiguity as to who should press which key and when (Novembre, Ticini, Schütz-Bosbach, & Keller, 2012). Therefore, there must be some degree of agent specificity for successful joint action to occur.

Two different models have been proposed to explain how a sense of agency is achieved. The first is a posdictive model, in which a sense of agency can only be established after an action has been performed. Wegner's theory of apparent mental causation (Wegner & Sparrow, 2007; Wegner, Sparrow, & Winerman, 2004; Wegner & Wheatley, 1999) posits that people experience agency based on three factors: (i) if thoughts of an action occurred immediately prior to the observation of the action (priority), (ii) if the observed action is consistent with one's intended action (consistency), and (iii) if the action cannot be accounted for by an alternative cause (exclusivity). Support for this account come from studies showing that people erroneously experience agency over actions they have not performed if the conditions of priority, consistency, and exclusivity had been met (Wegner et al., 2004; Wegner & Wheatley, 1999). Also, studies have shown that people are more likely to attribute their

own actions to others if the other people pose a plausible alternative cause (Wegner, Fuller, & Sparrow, 2003).

The second account of the experience of agency and action control is a forward model. It proposes that a sense of agency occurs based on the quality of the match between the predicted sensory consequences of a planned action and the actual sensory consequences of an executed action (Blakemore & Frith, 2003; C. D. Frith, Blakemore, & Wolpert, 2000). The better the match between the predicted and experienced action, the greater the sense of agency. Support for this forward model comes from studies showing that artificially creating discrepancy between the predicted and actual sensory consequences of a self-generated action modulates the sense of self-agency (Knoblich & Kircher, 2004; Sato, 2009; Sato & Yasuda, 2005). For example, Sato and Yasuda (2005) introduced temporal inconsistencies between a participant's button press and the playing of a tone. A sense of self-agency was reduced when a self-generated tone was unpredictable in its timing. However, an illusionary sense of agency was elicited when an externally generated tone happened to match the prediction made by forward model.

Novembre et al. (2012) employed a musical joint action task to investigate how the brain distinguishes between self- and other- generated actions during shared tasks. In their novel paradigm, pianists performed only the right-hand part of a piece of music they had previously learned to play bimanually. Participants were either led to (falsely) believe that the complementary left-hand part was being played by a co-performer (joint condition), or the left hand part was not played at all (self condition). This experimental manipulation served to induce either a mental representation of another individual playing the left-hand part, or the self playing the left-hand part (as they would when playing bimanually).

Single-pulse transcranial magnetic stimulation (TMS) was applied to the right primary motor cortex of the participants as they played. Motor-evoked potentials (MEPs) were recorded from the resting left forearm. The data revealed that MEP amplitudes associated with the self were lower than MEP amplitudes associated with the other's performance. Importantly, this result remained unchanged when the participant

could not directly see or hear the co-performer's actions but could infer the co-performer's presence. The authors interpreted their data as evidence that self- and other-related motor representations are not equivalent (however, it should be noted that although MEP amplitude is not a direct measure of cortical inhibition, it is possible that this pattern of results may reflect the inhibition of any movements of the left hand in the self condition). Novembre et al.'s (2012) findings suggest that musicians form agent-specific motor representations when performing joint musical tasks. This was true even when the music created by the co-performer was muted, suggesting that agent-specific representations occur in response to the potential for interaction. The authors proposed that while the representation of self- and other- generated actions recruit the same general neuroanatomical structures within the motor system, the neurophysiological representations are not fully equivalent within the interactive brain

Summary: Many tasks are performed in conjunction with other individuals. The coordination of behaviours can either be planned or emergent. There is evidence to suggest that executed and observed actions are cognitively represented in a functionally equivalent way, prompting the question of how self- and other- generated actions can be distinguished to create a sense of agency. One theory is that agency is determined after the event, by evaluating whether an observed action was consistent with a previously formed motor plan, or could be attributed to an external source. Alternatively, the forward-model proposed that agency was determined as soon as an observed action was perceived as being consistent with the predicted perceptual consequences of one's action intentions. Recent evidence from a musical paradigm found evidence of different cortical activation when representing the identical actions generated by the self and other. Therefore, while motor representations of the self and other recruit the same neuroanatomical resources, their neurophysiological representations are not fully equivalent.

2.8 Task Sharing and Task Co-representation

The previous section discussed how interpersonal co-ordination can occur emergently through direct action-perception couplings or be planned ahead of time. One means of predicting another individual's behaviour is to know the stimulus conditions under which they are required to perform a particular action (Sebanz, Bekkering, & Knoblich, 2006b; Sebanz & Frith, 2004). For example, when crossing a busy road, a red traffic signal allows a pedestrian to predict that cars are likely to stop. The ability to predict that it should be safe to cross the road when the traffic signal is red requires the participant to have 'co-represented' their own stimulus-response rules (i.e., when the car has stopped, cross the road) and the drivers' stimulus-response rules (i.e., when the traffic light is red, stop the car).

Evidence for task co-representation comes from a number of different paradigms. For example, in a study by Kilner et al.(2004) participants observed the conditions under which a co-actor consistently performed particular actions. For instance, when an object was presented in a specific colour they would observe an actor perform a grabbing motion, otherwise the co-actor's hand would remain still. Brain imaging data (based on EEG) revealed enhanced motor activity in the dorsal premotor cortex originating after the participant observed a cue for the co-actor's actions but *before* the action was performed. Kilner and colleagues proposed that knowledge of a co-actor's upcoming action is sufficient to excite one's own motor system, allowing individuals to anticipate (rather than simply react to) the actions of others'. In another study, van Schie et al. (2004) found that when participants were aware of an actor's stimulus-response rules, observing the actor perform an incorrect action elicited medial frontal activity similar to that elicited by making an error oneself.

Importantly, there is a growing body of evidence supporting that proposal that co-actors represent each other's tasks irrespective of whether doing so provides any benefit to the self (Atmaca, Sebanz, & Knoblich, 2011; Böckler, Knoblich, & Sebanz, 2012; Milward, Kita, & Apperly, 2014; Sebanz, Knoblich, & Prinz, 2003, 2005). This

includes complete 'division of labour' scenarios where the task components can be performed independently without any interpersonal coordination. The most popular paradigm used to investigate unwarranted task co-representation is the joint (or social) go/no-go Simon task (Sebanz et al., 2003). In the standard Simon task (Simon, 1990; Simon & Rudell, 1967), a single participant carries out a spatially defined response to a non-spatial attribute of a stimulus (e.g., press the button on the left if the stimulus is red, press the button on the right if the stimulus is blue) that appeared either to the left or right of a central fixation mark. Despite the fact that stimulus location is task-irrelevant, participants are typically faster to perform a correct response when the response button spatially corresponds with the stimulus (i.e., both on the left or both on the right). This stimulus-response compatibility effect is known as the Simon Effect (see Hommel, 2011 for a review).

Importantly, when a single participant is only required to respond to one of the two stimuli by pressing only one of the two response buttons, rendering the task a go-nogo task (i.e., press the left button if the stimulus is blue, if the stimulus is red refrain from responding), the Simon effect is not observed (Hommel, 1996). However, if a co-actor performs the complementary go-nogo task alongside the participant (i.e., the co-actor presses the button on the right if the stimulus is red) the Simon effect is re-established across the dyad (Sebanz et al., 2003). This 'joint Simon effect' (also referred to as the social Simon effect) has been explained in terms of co-actors automatically co-representing each other's stimulus-response rules (Sebanz et al., 2003). The amalgamated shared task representation is therefore functionally equivalent to a cognitive representation of the solo standard Simon task (Knoblich & Sebanz, 2006; Sebanz et al., 2003, 2005).

However, subsequent studies have cast serious doubt on the social nature of the joint Simon effect (Stenzel et al., 2014). Most notably, a joint Simon like effect can be observed when the co-actor is replaced by an event-producing object such as a metronome or Japanese waving cat (Dolk et al., 2011; Dolk, Hommel, Prinz, & Liepelt, 2013). Accordingly, non-social accounts of the joint Simon effect have been proposed (Dolk et al., 2013; Philipp & Prinz, 2010; Prinz, 2015; Wenke et al., 2011). For example,

the 'referential coding account' suggests that participants encoded the location of their response button to help discriminate between self-generated and externally-generated events. The spatial coding of the participant's own response re-introduces spatial stimulus-response compatibility effects that are responsible for driving the Simon effect (Hommel, Müsseler, Aschersleben, & Prinz, 2001; Kornblum, Hasbroucq, & Osman, 1990; Liepelt, Strobach, Frensch, & Schubert, 2011; Sellaro, Treccani, Rubichi, & Cubelli, 2013). Different accounts of the Joint Simon effect will be discussed in greater detail in Chapter 7.

Despite the dubious social nature of the joint Simon effect, there are a number of paradigms demonstrating automatic task co-representation that do not rely on stimulus-response compatibility. For example, in a novel paradigm created by Knoblich and Sebanz (2005), pairs of participants responded independently to different features of the same stimulus. The stimulus was an image of a hand with a pointing index finger and on the finger was a coloured ring. One participant was required to perform a go-no/go response based on the direction the hand was pointed (i.e., press a button if the hand is pointing towards the right side of the display, withhold from responding if the hand is pointed towards the left side of the display), the other participant was required to make a go-nogo response based on the colour of the ring (i.e., press a button if the ring is red, withhold from responding if the ring is green). When the co-actors were aware of each other's stimulus-response rules, responding was significantly slower when the stimulus required a response from both actors (i.e., a hand pointing right wearing a red ring). This response interference even occurred when the co-actors could not observe each other's actions.

Sebanz and colleagues proposed that each actor had integrated their co-actor's action alternative within his or her action planning. Therefore, an action selection conflict arose when a stimulus required an action from both actors. This is because the participant had to prioritise the motor representation of their own action over the motor representation of the co-actor's action. Furthermore, ERP measurements recorded during no-go trials indicated increased response inhibition when the co-actor was required to make a response, compared to when the co-actor was required to refrain

from responding (Sebanz, Knoblich, Prinz, & Wascher, 2006). This response inhibition suggests that when a stimulus is observed it triggers a motor representation of the participant's own response and the response of the co-actor. In order to avoid responding out of turn, participants must prioritise task-relevant motor representations and inhibit task-irrelevant motor representations. Taken together, these findings suggest that: (i) the participant knew what their co-actor should do (i.e., when to perform a response), and (ii) the co-actor's task was represented in a functionally equivalent way to their own task.

A further study has indicated that co-actors automatically take in to account each other's focus of attention when performing independent tasks (Böckler et al., 2012). Pairs of participants were presented with a global figure (e.g., a large letter H) made up of local features (i.e., multiple small letter Es). When asked to respond to the identity of either the local or global feature, participants were slower to do so if a co-actor was required to have a different focus of attention from their own (i.e., the participant responds to the global figure while the co-actor responds to the local features, or vice versa). This finding was taken to suggest that the co-actor's attentional focus induced a conflict that influenced the selection of one's own attentional focus.

Of note, a number of studies have found evidence to suggest that 'task co-representation' can occur even in the absence of an intentional co-actor (Dolk et al., 2013; Dolk, Hommel, Prinz, & Liepelt, 2014; Vlainic, Liepelt, Colzato, Prinz, & Hommel, 2010). For example, Saunders, Melcher, and van Zoest (2017) asked pairs of participants to perform independent Stroop tasks (Stroop, 1935). The participants each pressed one of two buttons to indicate which of their two assigned target colours was present in a stimulus. Stimuli consisted of individual words presented in different colours. Participants were instructed to ignore the meaning of the text and respond only to pixel colour. Three types of colour-incongruent distractor words were used for the text: (i) names of colours from their own response set, (ii) names of colours from the co-actor's response set, and (iii) neutral words for colours not used as pixel colours. When participants performed the task in dyads, distractor words taken from the co-actor's response set caused greater response interference that the neutral colour words

and caused just as much interference as distractor words from their own response set. However, the elevated interference from the co-actor's response set remained even if the co-actor was absent or did not exist. This pattern of results suggests that awareness of an alternative response set can cause task interference even if it is not affiliated with an intentional co-actor.

The nature and limits of task co-representation are still poorly understood. Most claims over the automaticity of task co-representation are restricted to direct stimulus-response mappings (Sebanz et al., 2005). Furthermore, it is not always clear what elements of the task are being co-represented and whether task interference originates during stimulus processing or action selection (Wenke et al., 2011). Both of these matters will be addressed in greater depth in Chapter 7. However, it is tempting to speculate that task co-representation (automatic or otherwise) is a cornerstone of social cognition (Sebanz, Bekkering, et al., 2006b). The ability to represent another individual's task allows people to anticipate, rather than simply react to, each other's actions. What is apparent is that representations of one's own task set is not fully prioritised over an alternative task set, even when it would be beneficial to the self. Why, how and when co-representation occurs is still very much a live topic of research.

Summary: One way in which people can anticipate the behaviours of others is to learn the environmental conditions in which they perform specific actions. There is evidence to suggest that people quasi-automatically represent the stimulus-response rules of nearby actors, even if doing so provides no benefit to their own task performance. It has been proposed that task co-representation is triggered by the social environment. However, awareness of an alternative task set can cause interference even if it is not associated with a present co-actor. Many questions remain as to the social nature of task co-representation.

2.9 Chapter Overview

Successful social interaction requires an ability to cognitively represent the actions, emotions, perspectives, and beliefs of other individuals. Importantly, there is compelling evidence to suggest that the neural architecture used in representing others is the same as that used to represent the self. For instance, observing another individual perform an action activates a corresponding motor representation of that action in the observer. The ability to distinguish between self- and other- representations and manipulate the relative activation of each, is crucial for successful social interactions. For instance, when taking the perspective of another individual it is imperative that the representation of one's own perspective is temporarily suppressed.

Neuroimaging and neurostimulation data have implicated two key regions of the brain in self-other control: the TPJ, which contributes to imitation control, theory of mind, and perspective taking; and the supramarginal gyrus, which contributes to the control of empathy. The disruption of self-other control has been associated with socio-cognitive disorders such as autism spectrum disorder, mirror-touch synaesthesia, and schizophrenia.

Furthermore, self-other control is important during task sharing. The successful coordination of behaviour can occur emergently (as a result of direct action-perception links) or be pre-planned (via the co-representation of a co-actor's task). A recent line of research suggests that people have an automatic tendency to represent the task rules of nearby others, even when interpersonal coordination is neither required nor beneficial to one's own task performance. Therefore, when performing a task within a shared environment self-other control is required to prioritise the representation of one's own task over the representation of a co-actor's task.

3 RESEARCH OUTLINE

In Chapter One and Chapter Two I provided an overview of evidence for two key claims in social psychology:

1) **Human cognition is fundamentally self-biased**. The way in which we perceive the world is shaped by the drive for self-preservation and self-enhancement. Accordingly, information relevant to the self is prioritised for cognitive processing over information associated with other individuals.
2) **Human cognition has evolved to understand others**. Humans are highly social animals and social-interaction is an important aspect of daily life. The ability to identify with others and take their perspective plays a fundamental role in intersubjective transactions. In order to successfully interact with conspecifics, individuals must cognitively represent the mental states and actions of others in a way that is relatable to their own.

These two statements appear to contradict one another. However, the wealth of evidence for both suggests that there must exist a mechanism that controls the prioritisation of self- and other-related representations during social interaction. The overall aim of my thesis was to expand current understanding of self-other control processes across different social and non-social contexts. I addressed three core research questions:

Research question 1: Are self-associated stimuli prioritised for processing over stimuli associated with a co-actor?

As is the standard for psychology experiments, the majority of tasks investigating self-biases have been performed by individual participants in well controlled laboratory settings. Accordingly, participants typically responded to self- and other associated

stimuli in social isolation. It is possible that it is only within non-social environments that people allow their egocentricity to go unchecked.

In daily life many tasks are performed either in conjunction with, or in the presence of, other individuals. Even when you do not intend for social-interaction to occur, it is often beneficial to be aware of what nearby others are doing. For example, when out jogging in the park it makes sense to anticipate whether a nearby game of catch could result in an errant ball flying in your direction. In other scenarios, such as working in a shared office, there may be a greater incentive to ignore other people in order to focus on the task at hand. Therefore, it is possible that the strength of the self-prioritisation effect is modulated by the social context in which a task in performed.

Most scientific investigations of self-prioritisation have adopted Sui et al's (2012) perceptual matching paradigm. In this relatively simple task participants associate three geometric shapes with themselves (i.e., I am a triangle), a familiar other (i.e., my mother is a square) and an unfamiliar other (i.e., a stranger is a circle). Following this,participants make speeded responses as to whether a presented shape has been paired with a matching social label (i.e., triangle - self) or a non-matching social label (i.e., triangle - stranger). Perceptual matching studies typically report a robust self-prioritisation effect; participants are significantly faster to match shapes to the self than to others. However, in these studies the 'others' have never been present in the task environment. In **Chapter 4** I present two studies investigating whether people prioritise processing of stimuli associated with the self over stimuli associated with an absent (**Experiment 1**) or present (**Experiment 2**) co-actor.

Research Question 2: Are stimuli associated with the self and others prioritised on the basis on likability?

Although self-biases are ubiquitous in cognition, the underlying mechanisms are unclear. While it is possible that the self is given special status within the brain, it is also possible that self-bias is a manifestation of a more general cognitive bias. For instance, it might be that self-relevant stimuli are prioritised for processing due to their inherent

reward value or positive affective valence. Studies that have attempted to discern whether the self-bias is equivalent to a positivity or reward bias have found that participant's sensitivity to the reward value or emotional valence of stimuli do not predict the participant's sensitivity to stimuli based on relevance to the self. So far, these studies suggest that self-bias is distinct from other cognitive biases.

Rather than taking a correlational approach, in **Chapter 5** I present two novel perceptual matching experiments that directly examine whether social stimuli are prioritised for processing based on their 'likability'. In **Experiment 3** participants associated geometric shapes with the self, a hero character from popular culture (Harry Potter) and a villain character from popular culture (Lord Voldemort). Participants then judged whether shape-name or shape-face pairings were matching or mismatching in accordance with the learned association. In **Experiment 4** participants performed the same perceptual matching task but with shapes associated with the self and two real-world political adversaries: President Barack Obama and President Donald Trump. I was interested in examining whether the subjective favourability of the 'other' individuals would influence the perceptual salience of the associated shape stimuli. The aims of the study were threefold:

1. To determine whether stimuli associated with a liked other are prioritised over stimuli associated with a disliked other.
2. To determine whether the self-prioritisation effect is influenced by the likability of the 'others'.
3. To determine whether the speed with which participants respond that the self-associated shape had been paired with a *non-matching* social label depends on the likeability of the label referent.

Chapter 6 continues this line of research by asking participants to choose their own 'favourable' or 'unfavourable' politician to use as stimuli. In two experiments participants responded to shape stimuli associated with the self, a stranger, and either a politician they held a favourable opinion of **(Experiment 5)** or a politician they held an unfavourable opinion of **(Experiment 6)**.

Research Question 3: To what extent are people able to prioritise their own task over the task of nearby others?

There is evidence to suggest that having an awareness of the task rules of nearby others can interfere with own-task performance, even if interpersonal coordination is neither required nor beneficial. However, important questions remain unanswered about what exactly is being 'co-represented' when people perform independent tasks in close proximity to one another.

In **Chapter 7** I present a novel paradigm developed to specifically investigate whether co-actors truly represent each other's stimulus-response rules. In this task, pairs of participants were presented with a multifeatured stimulus. Each participant was required to make an *independent* binary-choice response depending on whether or not the stimulus contained their individually assigned set of target features. The stimulus had the potential to contain all, some or none of the co-actors' target features, and the presence or absence of one target feature had no relationship with the presence or absence of another. Therefore, the co-actor's tasks were completely independent from one another. The aim of the study was to determine whether a participant's response to the presence or absence of their own target features was sensitive to the presence or absence of the co-actor's target features. In **Experiment 7** the co-actors performed the task side-by-side and were fully aware of their partner's task rules. In **Experiment 8** the co-actors performed their tasks side-by-side but were unaware of each other's task rules (control experiment)**.** In **Experiment 9** the co-actors performed their independent tasks in separate rooms but were aware of each other's task instructions.

4 Self-bias in Shared versus Non-Shared Task Environments

4.1 Abstract

Tasks used to examine cognitive self-biases typically involve participants responding to self- and other- associated stimuli in social isolation. The work described in this chapter examined whether stimuli associated with the self are prioritised for perceptual processing over stimuli associated with a present other. Participants matched arbitrary shape stimuli to labels representing the self, a hypothetical stranger, and either a present (Experiment 1) or absent (Experiment 2) task partner. Participants were significantly faster to match stimuli to the self than to the partner, irrespective of whether the partner was present or absent. Furthermore, participants were no faster to match stimuli to a present task partner than to a hypothetical stranger. Cross-experiment comparisons revealed that responses were generally faster when the task was performed in social isolation (Experiment 1). However, self-match RTs remained stable across experiments. Responses to non-matching shape-label pairings were slower if the self-shape was paired with a non-matching label referring to the present partner than the stranger. No such label effect was found when the partner was absent. This was taken to indicate greater representational overlap between the self and the partner when the partner was physically present. Overall, there was no evidence that stimuli associated with a present individual (self or other) were prioritised for processing over information relating to an absent or hypothetical individual.

4.2 Background

Using a simple shape-label matching paradigm, Sui et al. (2012) demonstrated the striking phenomenon that assigning self-relevance to a visual stimulus biases perceptual processing. In their novel task, participants associated three geometric shapes with the self, a familiar other (i.e., a best friend) and an unfamiliar other (i.e., a hypothetical stranger). Participants were significantly faster to indicate that a shape was paired with a matching social label if the shape was associated with the self than if it was associated with a familiar or an unfamiliar other. This self-prioritisation effect was particularly compelling as, unlike face or name stimuli, the self-relevant stimuli were not over-learned.

In studies adopting the perceptual matching task to investigate the effect of self-relevance on perceptual salience, participants have typically responded to self- and other- associated stimuli in social isolation (Moradi et al., 2015; Stolte et al., 2017; Sui & Humphreys, 2015d, 2015b). Within the context of these studies there is a fundamental difference between 'Self' and 'Other' that has gone relatively unacknowledged: while the self is physically present in the environment (in the form of the participant) the other individuals are not. It is possible that stimuli associated with an individual who is directly perceivable (e.g., in the same room) will be prioritised for processing over stimuli associated with an individual who is absent or imagined. Although it is difficult to remove the self from the task, it is possible to add other people to the task environment. The aim of the present work was to investigate whether visual stimuli associated with the self are prioritised for perceptual processing over stimuli associated with present and absent others.

There are a number of reasons to predict that stimuli associated with a present individual will be prioritised for processing over information relating to someone who is absent or imagined. For one, there is an adaptive advantage to the efficient processing of information relating to someone who can effect change in the immediate environment (C. D. Frith & Frith, 2007; Mayer, Vuong, & Thornton, 2017; Tomasello,

1995). Individuals close in proximity to the self are a potential source of interaction, welcomed or otherwise. As such, humans have evolved to be highly tuned to signals from the environment that could potentially help them to understand, anticipate, and respond appropriately to the behaviours of nearby others (C. D. Frith & Frith, 2006a, 2007). Such signals can be communicated consciously or subconsciously between conspecifics, and include body movements (Constable, Pratt, Gozli, & Welsh, 2015; C. D. Frith & Frith, 2006a), facial expressions (Vuilleumier & Pourtois, 2007), eye gaze (Huang, Andrist, Sauppé, & Mutlu, 2015), and direct verbal communication (Garrod & Pickering, 2004). However, people have also learned to predict the behaviours of nearby others by understanding the environmental conditions under which they perform particular actions (Sebanz, Bekkering, et al., 2006b). For example, when crossing a busy road at a traffic signal, pedestrians understand that drivers are obligated to stop when the light directed at their lane turns red. There are many such situations in which paying attention to how information relates to nearby others can be indirectly relevant to the self.

Of note, research investigating task performance in shared environments has found that individuals pay attention to information relevant to the tasks of nearby others even when interpersonal co-ordination is not required (Atmaca et al., 2011; Kiernan et al., 2012; Sebanz, Bekkering, et al., 2006b; Sebanz et al., 2003, 2005). For example, when pairs of participants performed complementary reaction time tasks side-by-side, with each participant responding to a different dimension of a mutually attended stimulus (i.e., one responding to the colour of the stimulus and the other responding to the location), own-task performance was influenced by how the co-actor should be responding (Sebanz et al., 2005). This task interference occurred even when co-actors were unable to observe each other's actions. Therefore, task interference occurred because participants had co-represented each other's stimulus-response rules and were not always successful at supressing the representation of their co-actor's task in favour of their own. This is an indication that self-prioritisation effects may be weaker when there is information in the environment associated with a present co-actor.

4.2.1 Concreteness of shape-person associations

Unrelated to the social significance of information, it is possible that associations formed between a visual stimulus and an individual who is present in the room benefit from being more concrete than associations formed between a visual stimulus and an individual who is absent (i.e. imagined). In the field of psycholinguistics it is well-established that words representing concepts that are highly concrete and imageable (e.g., 'cat' or 'umbrella') are subject to faster processing (Bleasdale, 1987), better recall (Fliessbach, Weis, Klaver, Elger, & Weber, 2006; M. Hamilton & Rajaram, 2001a; L. M. Miller & Roodenrys, 2009), and superior lexical decision making (Schwanenflugel, Harnishfeger, & Stowe, 1988), than words that represent relatively abstract concepts (e.g., 'time' or 'thought'). Following this, stimuli used to represent an individual who can be directly perceived are inherently more concrete than stimuli used to represent an imagined individual.

A recent study conducted by Wade and Vickery (2017) found that the purported self-association advantage in perceptual matching might be confounded by the concreteness of the social labels. Their data revealed that concrete non-social labels such as 'Frog' and 'Snake' resulted in a processing advantage comparable to the label 'Self', relative to the low-concrete control label 'Other'. Therefore, high-concrete labels were prioritised over low-concrete labels, irrespective of self-relevance. However, it should be noted that across all their experiments, Wade and Vickery (2017) found a significant self-label advantage relative to all other labels (abstract or concrete, social or non-social); only the size of the self-advantage varied depending on the concreteness of the other labels. Therefore, the results of that work cannot rule out the possibility that self-relevance contributes to the label-matching results.

Of note, Wade and Vickery found that the self-label advantage was smallest relative to the label 'Greg', which referred specifically to the experimenter who gave instructions to the participants. Although Greg was not present in the room when the participants performed the perceptual matching task, a photograph of the experimenter

was presented on the display as they encoded the shape-label associations. It is possible that the label 'Greg' benefitted from the extra contextual and visual information available during encoding. This idea is consistent with Dual Coding Theory (Paivio, 1991), which proposes that words encoded using both verbal codes and visual codes are more accessible from memory than words encoded using verbal codes alone. Whereas representations of an absent or hypothetical individual are limited to whatever relevant information is stored in memory (verbal and/or visual), visual representations of present and observable individuals can be updated in real-time. Furthermore, the representation of that individual will be activated and primed whenever they are perceived, resulting in more efficient processing overall.

4.2.2 The Present Study

The present study was conducted to investigate the extent to which visual stimuli associated with the self are prioritised over stimuli associated with a physically present co-actor. Sui, He and Humphrey's (2012) perceptual matching study was adapted so that participants associated three shape stimuli (triangle, square, circle) with the social labels 'Self', 'Stranger', and 'Partner'. In Experiment 1 the label 'Partner' referred to a hypothetical task partner, but in Experiment 2 the same label referred to a present co-actor performing a perceptual matching task alongside the participant. Therefore, by comparing the results of the two experiments, we can determine whether: (i) visual stimuli associated with a physically present individual are subject to enhanced perceptual processing relative to visual stimuli that are associated with a physically absent individual, and (ii) whether visual stimuli associated with the self are prioritised over visual stimuli that are associated with a co-actor who is physically present.

4.3 Experiment 1: Solo Task (Self versus Partner and Stranger Associations)

In Experiment 1 participants performed a perceptual matching task in social isolation (i.e., there was no other person present in the task environment). Participants associated one shape with the self, one shape with a hypothetical stranger, and one shape with an imagined task partner. Participants then judged whether a randomly paired shape and label stimulus were correctly or incorrectly matched in accordance with the recently learned associations.

The purpose of Experiment 1 was threefold:

(i) To determine whether a self-prioritisation effect could be elicited using the aforementioned shape-label stimuli. That is, whether participants are faster and more accurate when matching a self-associated shape to an appropriate social label than when matching a partner- or stranger-associated shape to the respective labels.

(ii) To determine whether participants respond differently to the 'Stranger' and 'Partner' associated stimuli. As both the partner and the stranger were hypothetical individuals whom the participant had never met, it was predicted that stimuli associated with one would have no greater social salience than stimuli associated with the other. Any difference in the efficiency of responses may indicate that either the physical (e.g., word length) or the semantic properties (e.g., affective connotations) of the social labels caused a biased responding.

(iii) Most importantly to provide comparison data for Experiment 2, in which the label 'Partner' referred to a present co-actor, rather than an absent and hypothetical individual. The null hypothesis (H0) is that there will be no difference in RTs to shape stimuli associated with an absent partner (Experiment 1) and a present partner (Experiment 2). The alternative hypothesis (H1) is that participants will demonstrate a bias for stimuli associated with a present partner, similar to that of stimuli associated with the self. Therefore, it would be expected that participants would be more efficient at responding to stimuli associated with the present partner (Experiment 2) than the absent partner (Experiment 1).

4.3.1 Method

4.3.1.1 Participants

Eighteen students from the University of Warwick participated in the experiment (seven men; 18 to 28 years of age, M = 21.7). The sample size was decided based on previous perceptual matching studies investigating social salience (see Sui et al., 2012). Each participant gave informed consent and the study was approved by the Ethics Committees of University of Warwick. Participants received monetary compensation (£4) for their participation. All had normal or corrected-to-normal vision.

4.3.1.2 Stimuli and Apparatus

Displays were generated and responses recorded using custom written programs running on a PC attached to a 19" LCD monitor running at a resolution of 1440 × 900 and 60Hz screen refresh rate. All stimuli were presented in white (RGB values = 255,255,255) on a uniform grey background (RGB values = 128,128,128). On each trial one of three geometric shapes (triangle, square, or circle) each subtending 3.8° × 3.8° at a viewing distance of 57 cm, was presented above a central fixation cross (0.8° × 0.8°). In addition, a text label was presented below the fixation. Label stimuli consisted of the words "Self" (2.4° × 1°), "Partner" (4.5° × 1°), and "Stranger" (5.5° × 1°), presented in Arial font. The distance between the centre of the shape or the label stimulus and the fixation cross was 3.5°.

Participants made 'match' and 'non-match' responses on custom-made control panels. The control panels consisted of two horizontally aligned touch sensors embedded in a casing with dimensions of approximately 7cm × 7cm × 0.5cm. The touch panels were connected to an Arduino microprocessor board that was programmed to emulate a USB gamepad. The sensors produced no auditory feedback when triggered. Participants held the control panels with both hands, using the right

thumb to press the right sensory, and the left thumb to press the left sensor. The assignment of the left and right sensors to the match and non-match responses was counterbalanced across participants.

4.3.1.3 Procedure

In the first stage of this experiment participants were given verbal instructions by the experimenter to encode three geometric shapes (triangle, square, and circle) as the self, a hypothetical task partner, and a hypothetical stranger. For example, a participant could be told the following: "In this task you are represented by a triangle. Your task partner, who is performing the same task in another room, is represented by a circle. A stranger you have never met before is represented by a square." Participants had 60 seconds to encode the shape-person associations prior to the start of the perceptual matching task. No visual stimuli were presented on the display at this stage.

Immediately following the encoding stage, participants began the perceptual matching task. Each trial began with the presentation of a central fixation cross for 500 ms. Next, a pairing of a shape stimulus (triangle, circle, or square) and a label stimulus ('Self', 'Partner' or 'Stranger') was presented on the display for 100ms. The pairing either conformed to the just-learned associations (i.e., a matching pair), or were a recombination of a shape with a non-matching label (i.e., a non-matching pair). A blank display was then presented for a variable time ranging from 800 to 1,200 ms. During this time participants were expected to indicate whether the shape-label pair was matching or non-matching by triggering one of two response sensors. After the response window ended, feedback of 'correct response' or 'incorrect response' was presented at the centre of the display for 500ms.

Following 12 practice trials, each participant completed a total of 360 trials over four blocks. Within each block, each possible shape-label combination occurred equally often (10 trials per condition per block) in random order. As the stimuli consisted of three shapes and three labels, there were twice as many non-matching trials than

matching trials. Therefore, participants were required to perform a 'non-matching' response twice as often as a 'matching' response.

4.3.2 Results

Table 1 presents mean correct RTs and accuracy data for Experiment 1. Responses faster than 200ms and missed responses (> 1200 ms) were excluded from analysis (<1% of trials). Accuracy across conditions = .83 (mean proportion correct) and no speed-accuracy trade-offs were apparent.

Table 1. Mean correct RTs and Accuracy as a Function of Shape Association (Self, Partner, and Stranger) and Label (Self, Partner, and Stranger), and Response Type (Match, Mismatch) for Experiment 1.

Shape Association	Label	Response Type	RT	Proportion Correct
Self	Self	Match	771.53 (82)	.81 (.15)
	Partner	Mismatch	816.38 (95)	.87 (.12)
	Stranger	Mismatch	816.87 (88)	.93 (.07)
Partner	Partner	Match	828.36 (80)	.69 (.09)
	Self	Mismatch	790.73 (94)	.89 (.1)
	Stranger	Mismatch	830.52 (104)	.85 (.13)
Stranger	Stranger	Match	845.59 (102)	.70 (.12)
	Self	Mismatch	787.54 (82)	.87 (.12)
	Partner	Mismatch	831.98 (103)	.84 (.07)

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses.

4.3.2.1 Accuracy

A signal detection approach (d') was used as a measure of response accuracy. For each shape association (Self, Partner, Stranger) correct match responses were treated as 'hits', and correct non-match responses were 'correct rejections'. Figure 1 presents mean d' scores for each shape association, with a larger d' indicating greater sensitivity.

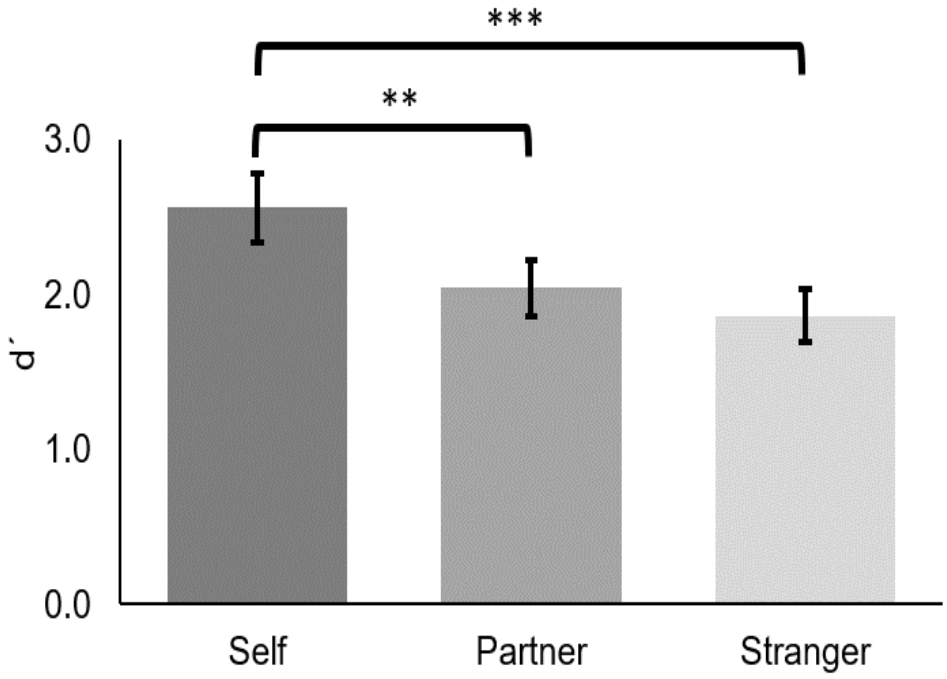

Figure 1. d' scores as a function of shape association for Experiment 1. Error bars represent the ± 1 standard error of the mean. Asterisks denote a significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$)

A within-subjects ANOVA revealed a significant main effect of shape association, $F(2, 34) = 9.19$, $p = .001$. The d' score was significantly larger for the self-associated shape than for either the partner-associated shape, $t(17) = 2.73$, $p = .01$, or the stranger-associated shape, $t(17) = 3.55$, $p = .002$. There was no significant difference in d' scores between the partner- and stranger- associated shapes, $t(17) = 1.7$, $p = .1$.

Therefore, participants demonstrated greater accuracy and sensitivity when responding to the self-associated shape than to either the partner- or stranger- associated shapes.

4.3.2.2 RT data

Correct RTs were analysed using a 3 (shape association: self vs. partner vs. stranger) × 2 (match condition: matching vs. non-matching) within-subjects ANOVA. This revealed no significant main effect of match condition, $F(1, 17) = .25$, $p = .62$. However, there was a significant main effect of shape association, $F(2, 34) = 695$, $p = .003$, which was qualified by a significant interaction with match condition, $F(2, 34) = 14.02$, $p < .001$. Accordingly, RT data for the matching and non-matching pairs are analysed separately below.

4.3.2.2.1 Matching trials

Figure 2a shows mean correct RTs for trials in which the shape and label were correctly paired. A one-way ANOVA revealed a significant main effect of shape association, $F(2, 34) = 11.02$, $p < .001$. Pairwise analyses revealed that match RTs were significantly shorter if the shape was associated with the self than with either the partner or the stranger; $t(17) = 3.37$, $p = .004$, and $t(17) = 3.67$, $p = .002$, respectively. There was no significant difference between match RTs for the partner- and stranger- associated shapes, $t(17) = 1.53$, $p = .14$.

4.3.2.2.2 Non-matching trials

Figure 2b shows the mean correct RTs for trials in which the shape and label did not match. A one-way ANOVA revealed no main effect of shape association, $F(2, 34) = 1.22$, $p = .31$, with all pairwise comparisons being non-significant ($t < 1.38$, ps $> .1$).

4.3.2.2.3 Specific label effects

On non-matching trials, a self-associated shape could form a non-matching pairing with either the partner label or with the stranger label (See Figure 2c). There was no significant difference in non-match RTs when the self-shape was paired with the partner label or the stranger label, $t(17) = .05$, $p = .96$. When the shape stimulus was associated with the stranger, non-match RTs were significantly faster if the shape was paired with a self-label than a stranger label, $t(17) = 4.06$, $p = .001$. On non-matching trails in which the shape was associated with the partner, RTs were significantly shorter for the self-label than the stranger-label, $t(17) = 3.41$, $p = .003$.

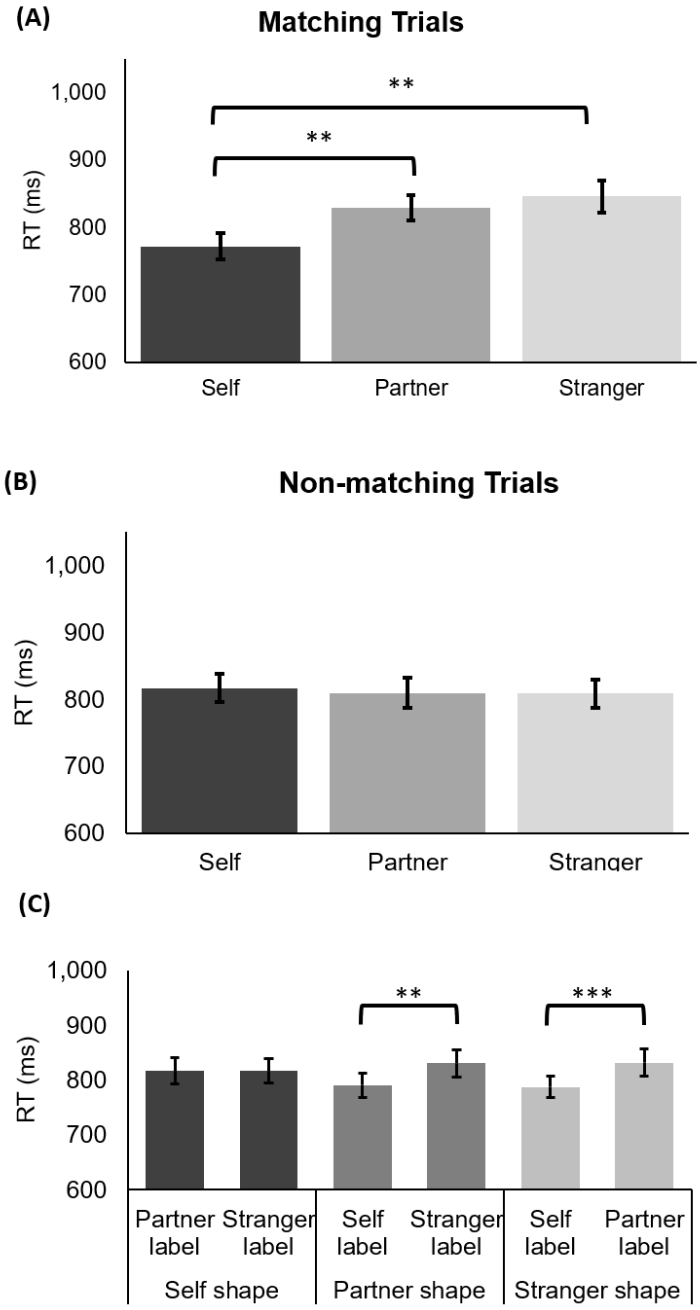

Figure 2. Mean correct RTs for matching (Panel A) and nonmatching (Panel B) shape-label pairs, as a function of shape association (Self, Partner, Stranger) for Experiment 1. Panel C shows mean correct RTs for each non-matching shape-label pairing. Error bars denote ± 1 standard error of the mean, asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

4.3.3 Discussion

A significant self-association advantage was found. Participants were both faster and more accurate when matching shape stimuli to the self, than when matching shape stimuli to a hypothetical stranger or task-partner. This self-prioritisation effect mirrors that first reported by Sui et al. (2012) as well as numerous subsequent studies (e.g., Moradi et al., 2015; Stolte et al., 2017; Sui & Humphreys, 2015d, 2015b).

Participants were no faster at matching appropriate labels to the stranger-associated shape than to the partner-associated shape. Thus, novel shape stimuli associated with the concept of a hypothetical 'Partner' appear to be of similar low salience to novel shape stimuli associated with the concept of a 'Stranger'. This is unsurprising, given that both the stranger and the partner were unknown hypothetical individuals. As participants had no clear visual representation of either the stranger or the participant, the associations formed between a shape and an abstract individual are likely to be less concrete than associations formed between a shape stimulus and the self (Czienskowski & Giljohann, 2002; Paivio, Yuille, & Madigan, 1968).

Furthermore, participants were no faster to respond that a self-associated shape had been paired with a non-matching label if the label was 'Partner' than when the label was 'Stranger'. Therefore, it is unlikely that differences in either the semantic or the physical properties of the partner and stranger labels influenced participant's responses. This finding is in line with initial control studies conducted by Sui et al. (2012), which found that word length, frequency and concreteness of the social labels had no effect on perceptual matching performance. However, participants were significantly faster to indicate that the partner- and stranger- associated shapes had been paired with a non-matching label if the label referred to the self. This label effect is further evidence of a self-association bias (Humphreys & Sui, 2015a; Sui & Humphreys, 2015c).

The results of the current experiment also provided no evidence to suggest that the emotional valence of the word 'Stranger', which often has negative connotations (e.g., 'stranger danger'), or the word 'Partner', which often has positive connotations

(e.g., a romantic partner), influenced perceptual matching. This finding meshes with that from another previous matching study conducted by Wade and Vickery (2017), in which participants were primed to believe that the social label 'Stranger' referred to a dangerous individual. Despite this additional task demand, no benefit was found for the stranger label relative to the self-label or the neutral label 'Other'. Therefore, it is unlikely that emotional valence of social stimuli is the key metric by which people prioritise perceptual processing.

It is possible that the performance benefit for the self-associated stimuli could reflect a true self-relevance effect, as was initially proposed by Sui et al. (2012). However, it remains to be determined whether the physical presence of the self in the task environment biases the processing of self-associated stimuli. While it is difficult to remove the self from the task, it is possible to add other people to the task environment. Therefore, the purpose of Experiment 2 was to determine the degree to which people prioritise self-relevant stimuli over other-relevant stimuli when performing a task in a shared environment.

4.4 Experiment 2: Self- versus physically present partner and hypothetical stranger associations

In Experiment 2 pairs of participants performed the same perceptual matching task side-by-side. Each participant associated one shape to the self, one shape with a hypothetical stranger, and one shape with their present co-participant (or task partner). Therefore, the goal of Experiment 2 was to determine whether: (i) stimuli associated with a present-other would be prioritised over stimuli associated with an absent-other, and (ii) whether stimuli associated with the self would be prioritised over stimuli associated with present-other.

4.4.1 Method

4.4.1.1 Participants

Eighteen students from the University of Warwick participated in the experiment (seven men; 18 to 24 years of age, M = 20.2). Participants received monetary compensation (£4) for their participation. All had normal or corrected-to-normal vision. Participants performed the task in pairs. Pairings were not matched for age or gender, however, all participants confirmed that they had not met their co-participants prior to the experiment.

4.4.1.2 Stimuli and Procedure

The stimuli and general task design of Experiment 2 were the same as that of Experiment 1. The key change was that instead of individual participants performing solo perceptual matching tasks in social isolation, pairs of participants performed the same perceptual matching task side-by-side (and in response to mutually attended stimuli). In this version of the task the label 'Partner' referred specifically to a physically present co-actor, rather than a hypothetical absent task partner.

Co-participants were asked to encode the exact same shape-label associations. For example, the following spoken instructions were directed at each participant individually (but within earshot of one another): "You [directed at each participant in turn] are represented by a triangle, your task partner is represented by a circle, and a hypothetical stranger is represented by a square". During the perceptual matching task both co-participants responded to the same stimuli, presented on the mutually attended monitor. They each independently indicated whether the presented shape-label pairings conformed to the recently learned shape-person associations by pressing one of two touch sensors on their personal control panels. The assignment of the touch sensors to indicate a matching or non-matching shape-label pairing was counterbalanced between but not within co-participants.

Each trial ended when the response window timed out (irrespective of whether either or both participants had made a response). Feedback of "correct response", "incorrect response" or "missed response" was presented on the left half of the display for the participant seated to the left, and the right half of the display for the participant seated on the right. Therefore, co-participants could only determine how their co-actor had responded at the end of the trial.

4.4.2 Results

Table 2 presents mean correct RTs and accuracy data for Experiment 2. Responses faster than 200ms and missed responses (> 1200 ms) were excluded from analysis (<1% of trials). Accuracy was generally high across conditions (mean proportion correct = .80) and no speed-accuracy trade-offs were apparent.

Table 2. Mean correct RTs and Accuracy as a Function of Shape Association (Self, Partner, and Stranger) and Label (Self, Partner, and Stranger), and Response Type (Match, Mismatch) for Experiment 2.

Shape Association	Label	Response Type	RT	Proportion Correct
Self	Self	Match	821.33 (86)	.91 (.08)
	Partner	Mismatch	919.43 (69)	.86 (.12)
	Stranger	Mismatch	875.01 (88)	.95 (.05)
Partner	Partner	Match	945.48 (70)	.80 (.13)
	Self	Mismatch	856.38 (78)	.93 (.07)
	Stranger	Mismatch	910.78 (67)	.91 (.07)
Stranger	Stranger	Match	953.33 (86)	.79 (.12)
	Self	Mismatch	842.03 (73)	.93 (.08)
	Partner	Mismatch	917.82 (74)	.90 (.08)

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses.

4.4.2.1 Accuracy

As with Experiment 1, signal detection (d') was used as a measure of response sensitivity. Figure 3 presents mean d' scores for each shape association, with a larger d' indicating greater sensitivity in responding.

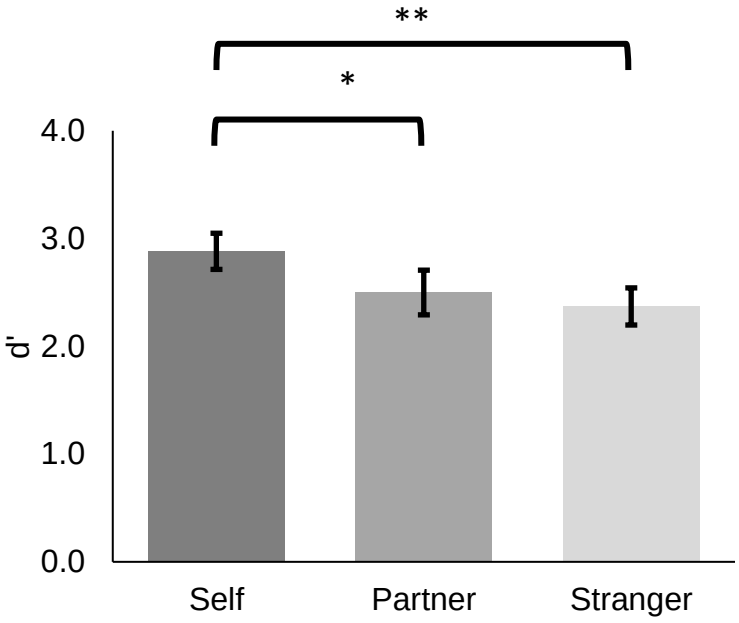

Figure 3. d' scores as a function of shape association for Experiment 2. Error bars denote ± 1 standard error of the mean. Asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$)

A within-subjects ANOVA revealed a significant main effect of shape association, $F(2, 34) = 4.96$, $p = .013$. d' was significantly larger for responses to the self-associated shape than to either the partner-associated shape or the stranger-associated shape, $t(17) = 2.3$, $p = .04$, and $t(17) = 2.88$, $p = .01$, respectively. There was no difference in d' scores for the partner and stranger shapes, $t(17) = .8$, $p = .44$. Therefore, participants demonstrated greater sensitivity when responding to the self-associated shape than when they responded to the partner- or stranger- associated shapes.

4.4.2.2 RT data

A 3 (shape association: self vs. partner vs. stranger) × 2 (match condition: matching vs. non-matching) within-subjects ANOVA was conducted on correct RTs. There was no main effect of match condition, $F(1, 17) = 3.58$, $p = .08$. However, there was a significant main effect of shape association, $F(2, 34) = 20.33$, $p < .001$, which was qualified by a significant interaction with match condition, $F(2, 34) = 45.41$, $p < .001$. Accordingly, RT data for the matching and non-matching pairs will be analysed separately below.

4.4.2.2.1 Matching trials

Figure 4a shows mean correct RTs for trials in which the shape and label stimuli were matching. A one-way ANOVA revealed a significant main effect of shape association, $F(2, 34) = 36.01$, $p < .001$. Match RTs were significantly faster when the shape was associated with the self than with either the partner or the stranger; $t(17) = 8.14$, $p < .001$, and $t(17) = 6.41$, $p < .001$, respectively. There was no significant difference in match RTs for the partner- and stranger- associated shapes, $t(17) = .61$, $p = .63$.

4.4.2.2.2 Non-matching trials.

Figure 4b shows the mean correct RTs for trials in which the shape and label stimuli did not match. A one-way ANOVA revealed no significant main effect of shape association, $F(2, 34) = 2.75$, $p = .08$.

4.4.2.2.3 Specific label effects

On non-matching trials each shape could be paired with one of two labels (See Figure 4c). When the shape was associated with the self, non-match RTs were significantly faster if the label referred to the stranger than to the partner, $t(17) = 4.68$, $p < .001$. When the shape stimulus was associated with the partner, RTs were significantly faster when the non-matching label referred to the self than the stranger, $t(17) = 4.11$, $p = .001$. When the shape stimulus was associated with the stranger, non-matching RTs were faster when the label referred to the self than the partner, $t(17) = 6$, $p < .001$.

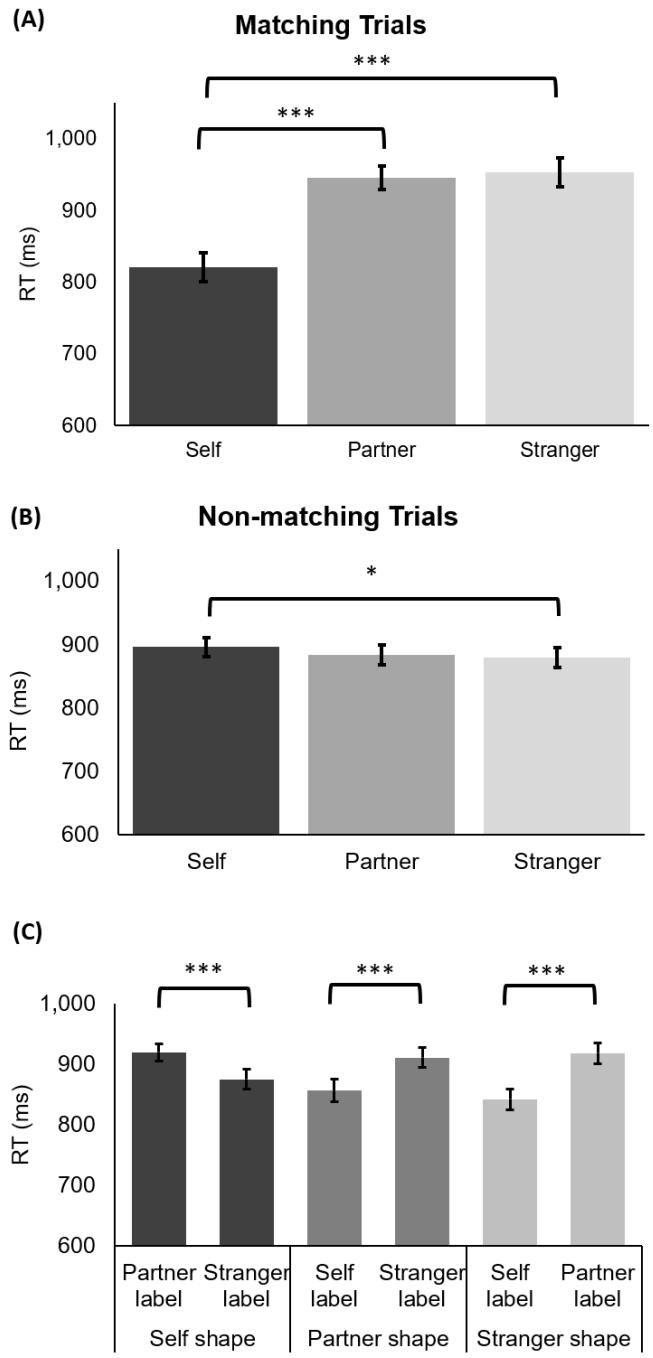

Figure 4. Mean correct RTs for matching (A) and nonmatching pairs (B) as a function of shape association (Self, Partner, Stranger) for Experiment 2. Panel C shows mean correct RTs for each non-matching shape-label pairing. Error bars denote ± 1 standard error of the mean, asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$)

4.4.3 Discussion

In Experiment 2 participants associated three arbitrary shape stimuli with the self, a hypothetical stranger, and a present task-partner. The aim of the experiment was to determine whether (i) self-associated stimuli would be prioritised for perceptual processing over stimuli associated with a present co-actor, and (ii) whether stimuli associated with a present co-actor would be prioritised over stimuli associated with an absent stranger. The data revealed a robust self-association benefit relative to both partner- and stranger- associations. This self-association advantage indicates that novel stimuli associated with the self have heightened social salience relative to similar stimuli associated with an individual who is also present in the task environment. Furthermore, no benefit was found for stimuli associated with a present-other over stimuli associated with an absent-other. Therefore, this experiment provides no evidence that 'tagging' an arbitrary shape to a physically present person results in a stronger or more salient association than tagging a shape to a physically absent person.

Of note, when a self-associated shape was paired with a non-matching social label, responses were significantly slower when the label referred to the present partner than to the absent stranger. This finding may reflect a greater overlap between representations of the self and a present other, than representations of the self and a hypothetical stranger. This possibility will be discussed further in the general discussion section of this chapter.

4.5 Cross Experiment Comparisons

Cross-experiment comparisons were conducted to determine whether the physical presence of the partner in the task environment influenced responses to the self, stranger, and partner associated stimuli.

4.5.1 Accuracy

An ANOVA with a within-subjects factor of shape association (self, partner, stranger) and a between-subjects factor of partner presence (present, absent) was conducted on d' scores. There was a significant main effect of shape association, $F(2,68) = 13.79$, $p < .001$, $\eta^2 = .29$, reflecting the robust self-association benefit found across both experiments. No main effect of experiment was found, $F(1,34) = 3.54$, $p = .07$, $\eta^2 = .09$, and there was no significant interaction between shape association and experiment, $F(2,68) = .33$, $p = .72$, $\eta^2 = .007$.

4.5.2 RT data

Analyses were conducted using mixed ANOVAs with shape association (self, partner, stranger) as a within-subjects variable and partner presence (present, absent) as a between-subjects variable. Figure 5 displays the mean match RTs for self-, partner-, and stranger- associated shapes as a function of partner presence. Figure 6 shows the equivalent non-match RTs. Match and non-match RTs will be analysed separately.

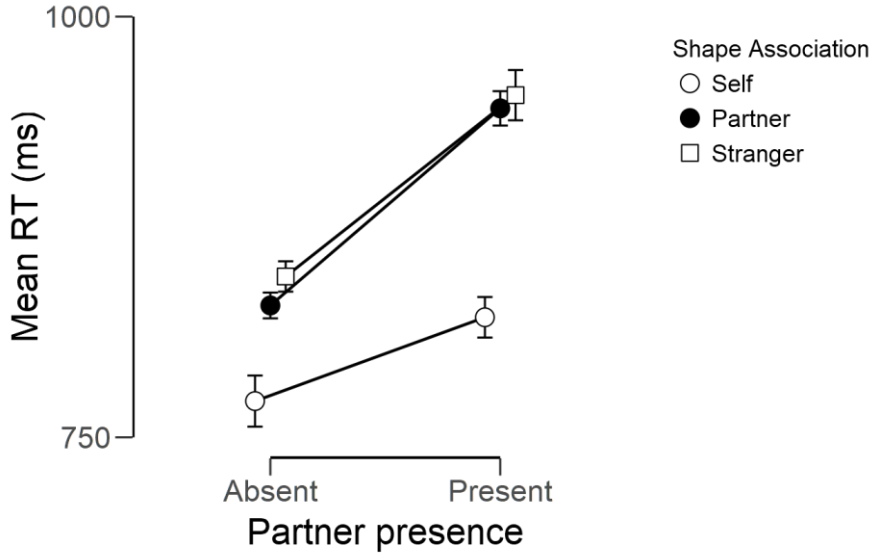

Figure 5. Mean correct RTs for match trials as a function of shape association (self, partner, stranger) for partner absent (Experiment 1) and partner present (Experiment 2) conditions. Error bars denote the standard error.

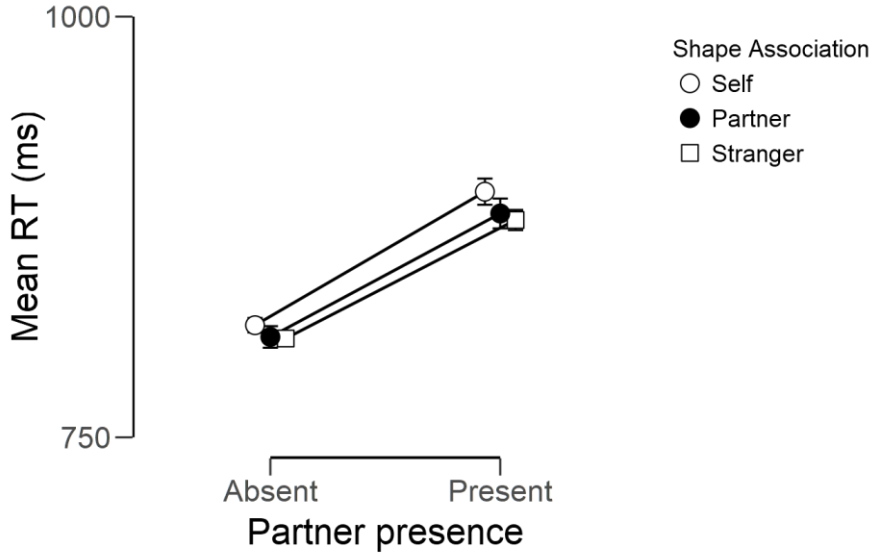

Figure 6. Mean correct RTs for non-match trials as a function of shape association (self, partner, stranger) for partner absent (Experiment 1) and partner present (Experiment 2) conditions. Error bars denote the standard error

4.5.3 Match RTs

There was a significant main effect of partner presence on match RTs, $F(1,34) = 13.75$, $p < .001$, $\eta^2 = .29$. Responses were significantly faster if the task was performed in social isolation (Experiment 1) than if the task was performed alongside the partner (Experiment 2). There was also a significant main effect of shape association, $F(2,68) = 43.81$, $p < .001$, $\eta^2 = .53$, which was qualified by a significant interaction with partner presence, $F(2,68) = 4.61$, $p = .01$, $\eta^2 = .06$. Pairwise analysis revealed that match RTs for the self-shape did not significantly differ across experiments, $t(34) = -1.78$, $p = .09$, $d = .59$. Therefore, self-match RTs were not influenced significantly by the presence of the partner in the task environment. In contrast, both partner-match and stranger-match responses were significantly slower if the partner was present in the room than if the task was performed in isolation; $t(34) = -4.67$, $p < .001$, $d = -1.55$ and $t(34) = -3.42$, $p = .002$, $d = -1.14$, respectively.

Stranger-match responses were then removed from the analysis in order to specifically investigate whether the presence of the partner influenced the size of the self-association advantage over partner-association. A significant interaction was found between partner presence (present, absent) and shape association (self, partner), $F(1,34) = 8.76$, $p = .006$, $\eta^2 = .08$. As previously established, this interaction was driven by a significant slowing of partner-match responses if the task was performed alongside the co-actor (Experiment 2) compared to if the task was performed in isolation (Experiment 1). No such change was found for self-match responses.

Previous analysis also found no significant difference between partner- and stranger match RTs in either Experiment 1 or Experiment 2 when analysed individually. Furthermore, both partner- and stranger- match responses were significantly slower if the partner was present in the room. A follow-up 2 (partner presence: present, absent) x 2 (shape association: partner, stranger) mixed ANOVA revealed no significant interaction, $F(1,34) = 0.23$, $p = .64$, $\eta^2 = .006$. Therefore, the presence of the partner in

the task environment slowed partner-match and stranger-match responses to a similar degree.

4.5.4 Non-match RTs

The presence of the partner in the task environment significantly slowed non-match responses, $F(1,34) = 8.11$, $p = .007$, $\eta^2 = .19$. There was also a main effect of shape association, $F(2,68) = 3.95$, $p = .02$, $\eta^2 = .1$, reflecting faster non-match responses when the shape was associated with the self than when it was associated with either the partner or the stranger; $t(34) = -1.98$, $p = .05$, $d = .33$ and $t(34) = -2.79$, $p = .008$, $d = .47$, respectively. No difference was found between non-match RTs when the shape stimulus was associated with the partner and the stranger, $t(34) = 0.54$, $p = .59$, $d = .09$. There was also no significant interaction between shape association and partner presence for non-match RTs, $F(2,68) = 0.48$, $p = 62.$, $\eta^2 = .01$.

4.5.5 Specific label effects on self-non-match RTs

Previous analysis revealed that when the partner was present (Experiment 2), participants were significantly slower to make a non-match response if the self-associated shape was presented with a label referring to the partner than the stranger. However, this label effect was not found when the partner was absent (Experiment 1). Confirming this difference, a cross-experiment analysis revealed that the interaction between partner presence (present, absent) and label type (partner, stranger) was significant, $F(1,34) = 11.77$, $p = .002$, $\eta^2 = .21$. Figure 7 shows the mean non-match RTs for the self-associated shape as a function of label type and partner presence.

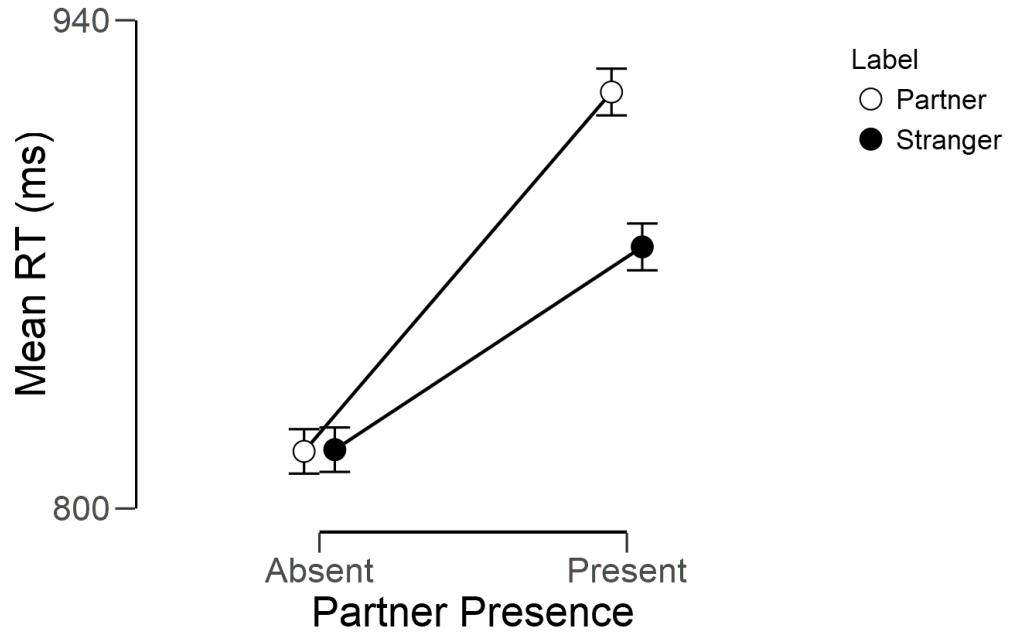

Figure 7. Mean non-match RTs for the self-associated shape as a function of label type and partner presence. Error bars denote standard errors.

4.6 General discussion

It has been well-established across a number of cognitive domains that stimuli associated with the self are subject to enhanced processing relative to stimuli associated with other individuals (see Chapter 1 for an overview). However, in most experiments investigating cognitive self-biases, participants are required to respond to self- and other- associated information in social isolation. Therefore, while the self always had a concrete physical presence within the task environment, the 'others' had to be imagined. It is possible that purported self-biases had been confounded by the ability to 'tag' information to a physical, concrete and perceptible individual. The purpose of the present study was to determine whether information associated with the self would be

prioritised over information associated with another individual who was also physically present in the task environment.

In two experiments, participants were required to associate three geometric shapes with labels referring to the self, a hypothetical stranger, and a task partner. In Experiment 1 the partner label referred to a hypothetical individual performing the same task in a different room. In Experiment 2 the partner label referred specifically to the (real) individual performing the same task alongside the participant. Once the shape-label associations had been encoded, participants were then presented with random shape-label pairings and were required to indicate whether the pairings conformed to the previously learned associations (by performing a match response) or were a recombination of a shape stimulus with a non-matching label (by performing a non-match response).

Consistent with the vast majority of perceptual matching studies which have investigated social salience effects (Humphreys & Sui, 2015b; Stolte et al., 2017; Sui et al., 2012; Sui & Humphreys, 2015d, 2015a; Wade & Vickery, 2017), a robust self-prioritisation effect was found. Participants were significantly faster to match shape stimuli to the self than to either the stranger or the task partner. Importantly, this was true irrespective of whether the task partner was present within, or absent from, the task environment. No difference was found in the efficiency with which participants matched stimuli to the hypothetical stranger and the hypothetical partner (Experiment 1) or the hypothetical stranger and the present task partner (Experiment 2). Overall, the study found no evidence to suggest that the ability to 'tag' visual stimuli to a physically present individual biases performance on a perceptual matching task. Furthermore, this indicates that the self-association bias found in perceptual matching is unlikely to be due to the physical presence of the self in the task environment.

The finding that information associated with a present-other was not prioritised for processing over information associated with an absent-other was somewhat surprising. It was predicted that stimuli associated with a present other would benefit from being more concrete. This prediction was based on the results of a study by Wade

and Vickery (2017), who found that the concreteness of the social labels biased performance in perceptual matching tasks. One possible reason why the physical presence of the partner did not result in the prioritisation of partner-associated stimuli, was that the stimuli had become disassociated with the co-actor. Outside of the context of the task the term 'partner' can have a number of different referents (e.g. a romantic partner or a business partner). It is possible that the shape intended to be specifically associated with the co-actor actually became associated with a different partner or a more general conceptualisation of a partner.

A further important finding was that despite both partner-match and stranger-match responses being significantly faster if the task was performed in social isolation than alongside a co-actor, no such difference was found for self-match responses. The finding that self-match RTs remained stable across experiments was somewhat surprising given that 'audience effects', the tendency for people to perform differently when in the presence of others than when alone, are well reported in social psychology (Feinberg & Aiello, 2006; A. F. de C. Hamilton & Lind, 2016; Sanders, 1981). For one, a present co-actor produces salient events that are likely to distract attention away from one's own task (e.g., Baron, 1986). Why, then, is the processing of self-associated stimuli immune to such effects? It is well established that information most relevant to the self is granted prioritised access to limited cognitive resources. For example, own-face and own-name stimuli have been found to be resistant to the attentional blink (Shapiro et al., 1997) and repetition blindness (Arnell et al., 1999). Self-referent stimuli have also been found to make more salient targets and more potent distractors compared to equivalent stimuli associated with other individuals (Gronau et al., 2003; Keyes & Dlugokencka, 2014; Salomon, Lim, Kannape, Llobera, & Blanke, 2013). The present findings suggest that the privileged status of the self in cognition allows self-relevant information to easily outcompete even highly salient distractors.

Of note, the analysis of non-match responses revealed that participants were slower to respond when the self-associated shape was paired with a label referencing the partner than the stranger. However, this label effect only occurred if the partner was present in the task environment. It is possible that the greater the degree of

representational overlap between self and other, the more difficult it is to discriminate between self- and other- associated stimuli. In support of this idea, behavioural and neuroimaging data has found evidence that people automatically represent nearby others in a manner that corresponds with their self-representation (Brass & Heyes, 2005; Decety & Sommerville, 2003a). The functional similarity between representations of self and other allows for an individual to relate to, and interact with, others. Furthermore, studies of task performance in shared environments has provided evidence that participants represent the actions and even task-rules of nearby others, irrespective of whether interpersonal coordination in likely (Sebanz et al., 2003). Accordingly, the ease at which an individual can distinguish themselves from others is influenced by physical distance.

In sum, the main finding of the present study was that novel stimuli associated with the self were prioritised for perceptual processing over novel stimuli associated with other individuals also present within the task environment. In other words, there was no evidence to suggest that the self-association bias in perceptual matching is confounded by the absence of 'others' relative to the presence of the self. Furthermore, stimuli associated with present others was not prioritised over information associated with absent others. Overall, the ability to 'tag' novel stimuli to a physical body was not found to influence salience.

5 HERO VERSUS VILLAIN: THE ROLE OF LIKABILITY IN SELF-PRIORITISATION

5.1 Abstract

Self-associated stimuli are subject to enhanced cognitive processing relative to stimuli associated with other individuals. In this chapter I investigated whether this bias reflects the heightened emotional value or 'likeability' of self-stimuli. Participants associated three shapes with the self, a liked other, and a disliked other, before judging whether a series of shape-label or shape-face pairings represented correctly or incorrectly matched pairs. In Experiment 3 the liked and disliked others were a hero/villain combination from popular culture (Harry Potter and Voldemort). In Experiment 4 participants provided favourability ratings for two political figures (Barack Obama and Donald Trump), that were used to determine which politician they liked the most/least. Across experiments participants were more efficient at matching shapes to the self than to either the liked or disliked other. Furthermore, participants were significantly faster to respond when the self-shape had been incorrectly paired with a face/label referring to the disliked other than to the liked other. I propose that participants perceive greater representational incongruence between the self and a disliked other compared to a liked other, aligning with research showing that cognition is motivated by a need for positive self-regard.

5.2 Background

Self-biases are ubiquitous in social cognition. For example, people show enhanced recall for information that is associated with the self (e.g., Cunningham, Turk, Macdonald, & Macrae, 2008; Klein & Kihlstrom, 1986; Rogers, Kuiper & Kirker, 1977), are faster to recognise and more accurate at making perceptual judgements about their own face than the faces of others (Keyes & Brady, 2010; Sui, Chechlacz, & Humphreys, 2012), and switch attention automatically on mention of their own name (i.e., the cocktail party effect; Moray, 1959). However, it remains unclear how self-relevance is recognized in the brain, and what cognitive mechanisms are involved in the prioritized processing of self-associated information. One question currently under investigation is whether there are specialized mechanisms involved in the prioritization of self-relevant information, or whether self-biases arise from the operation of more generalized cognitive processes. For example, it might be that processing benefits afforded to self-referent stimuli (such as one's own face and name) are due to familiarity effects or even the positive emotional response that such stimuli typically provoke.

In an influential study, Sui et al. (2012) demonstrated that simply 'tagging' novel arbitrary stimuli with self-relevance results in a processing advantage for those stimuli. Participants were asked to form associations between three shapes (i.e., triangle, square and circle) and the self, a familiar other (such as a close friend), and an unfamiliar other (i.e., a named stranger). Once these associations had been encoded in memory, participants were presented with random shape-label pairings (e.g. a triangle shape paired with the word friend) and were asked to judge whether the pairing was correctly or incorrectly matched, in accordance with the just-learned associations. For matched pairings, responses were significantly faster and more accurate when the shape was associated with the self, than when the shape was associated with either the familiar or unfamiliar other. Therefore, Sui et al., provided compelling evidence that self-association has a powerful top-down influence on relatively low-level perceptual processes, irrespective of stimulus familiarity.

The question remains as to whether the enhanced processing of self-relevant stimuli can be accounted for in terms of a positivity bias. That is, whether cognition is shaped by an inherent drive to 'like' the self, perhaps as a means of self-preservation

(the protection of oneself from harm or death) or self-enhancement (taking a tendentiously positive view of oneself to maintain positive self-esteem (see Sedikides & Gregg, 2008 for a review)). A diverse body of research has accumulated demonstrating that people perceive themselves through metaphorical 'rose-tinted glasses' (see Cunningham & Turk, 2017, for a review). For example, self-evaluations in autobiographical memory tend to be more positively skewed compared to evaluations about others (D'Argembeau & Van der Linden, 2008; Walker, Skowronski, & Thompson, 2003); non-depressed individuals are more efficient at processing and recalling positive information about the self but negative information about others (Kuiper & MacDonald, 1982); and people tend to remember events in a way that emphasizes their most favourable attributes (Alicke & Govorun, 2005; Sedikides & Gregg, 2003, 2008). This skewing of positive attributions also extends to objects that one owns, known as the mere ownership effect (Beggan, 1992), to the extent that people place a higher value on an object *once* they own it than before they own it (the endowment effect, (Kahneman et al., 1990; Knetsch, 1989).

A recent study by Stolte, Humphreys, Yankouskaya, and Sui, (2017) investigated the relationship between positive emotion and self-processing by contrasting participants' performance on two versions of Sui et al.'s (2012) associative matching task. In one session participants were asked to associate three shapes with the self, a friend, and a stranger, before judging whether subsequently presented shape-label pairings matched or mismatched. In another session, the same participants were asked to associate three new shapes with three faces, each exhibiting a different facial expression; happy, sad, or neutral. The data revealed an advantage of self-association over friend and stranger association, as well as a significant benefit for associations with a happy expression compared with sad and neutral expressions. However, no significant correlation was found between the effects of self-association and happy-expression at the level of the individual. This finding suggests that although the positive emotional valence of a stimulus has a similar effect on perceptual salience as self-relevance, the mechanisms that drive these biases are not the same.

Similar experiments have been conducted to examine whether self-bias is equivalent to a reward bias (Sui et al., 2012; Sui & Humphreys, 2015d; Sui, Yankouskaya, & Humphreys, 2015). For example, Sui and Humphreys (2015) assigned either a large (£8) or a small (£2) reward value to the friend- and stranger- associated shapes, but no reward to the self-associated shape. Despite receiving no reward, participants were more efficient at responding to self-associated stimuli than to friend- and stranger- associated stimuli that had been assigned a low-reward value. However, there was no advantage of self-association relative to high-reward stimuli. This pattern of results suggests that self-relevance is processed in an equivalent manner to high-reward, however, in a similar way to positive valence effects, the strength of an individual's self-bias did not correlate with the strength of their reward bias. It is therefore likewise argued that self-bias and reward-bias have distinct origins.

Of note, Sui and Humphreys (2015b) also found that the relationship between self-bias and reward was influenced by the perceived personal distance between the self and the imagined stranger. Participants were asked to place marks on a straight line to indicate the personal distance between each of three individuals (self, friend, and stranger). The physical distance between the markings acted as an index of personal distance. For individuals who expressed a relatively close personal distance from strangers, greater sensitivity to reward predicted a reduction in the advantage of self-association over high-reward stimuli, indicating a common value system. In contrast, for participants who expressed a relatively large personal distance between the self and the stranger, no relation was found between their basic reward bias and the magnitude of their self-bias. Sui and Humphreys (2015b) proposed that the greater the personal distance between self and other, the more independent the self-bias is from other motivators of behaviour, such as reward.

There is also evidence to suggest that the self-bias extends to include group membership. Moradi et al. (2015) asked soccer fans to match geometric shapes to the logo of either: (i) the team they support (in-group), (ii) their main rival team (out-group rival), or (iii) a team that is not considered a close rival (out-group neutral). Participants showed greater efficiency and discriminability when responding to in-group stimuli,

relative to out-group stimuli (both rival and neutral). Of note, the in-group bias was found to correlate with measures of team satisfaction. However, no difference was found in the speed or sensitivity to which participants matched shapes with the rival and neutral teams. These findings suggest that the in-group bias was sensitive to the perceived emotional or reward value of being a member of the group, but responses to the out-group were not sensitive to the degree of rivalry or aversion felt towards the alternative team.

Another point of interest is whether a self-bias remains when other social stimuli convey threat. There are adaptive benefits to being able to attend, perceive, and respond quickly to information within the environment that is pertinent to the needs of the self. There seems no reason why this should not also extend to information relating to other individuals if it informs an observer about potential threats. A large number of visual detection studies have demonstrated that humans are more efficient at detecting a threatening stimulus compared to a wide variety of non-threatening stimuli (Eastwood, Smilek, & Merikle, 2001; Öhman, Lundqvist, & Esteves, 2001; L. J. Schmidt, Belopolsky, & Theeuwes, 2012, 2015). However, little research contrasting self-bias and threat-bias has been conducted.

Wade and Vickery (2017) investigated whether shapes tagged with a threatening label are subject to the same salience-enhancing mechanisms that shapes tagged with a self-referent label seemingly access. Using an adapted version of Sui et al's (2012) perceptual matching paradigm, participants were asked to associate shapes with labels referring to the self ('You'), a neutral other ('Other') and either a label referring to a dangerous other ('Stranger') or a dangerous animal ('Snake'). Their results revealed that the labels 'You' and 'Snake' were subject to prioritised processing relative to the label 'Other'. However, the label 'Stranger', was not. As threat was not found to be a reliable determinant of stimulus prioritisation, it is possible that other factors may be involved. For example, Wade and Vickery (2017) proposed that the oddball nature of the label 'Snake' may have biased responding. This theory was supported by the fact that the non-threatening label 'Frog' received a similar processing benefit, highlighting that social-salience effects can be confounded by non-social properties of the label.

In addition, it is possible that the varying degree of concreteness or imageability of the formed associations pose a serious confound when investigating social salience effects. Concreteness pertains to the extent to which a concept represents a perceptible entity (see Paivio, Walsh, & Bons, 1994). Research has shown that highly concrete words are subject to enhanced memory recall (Campoy, Castellà, Provencio, Hitch, & Baddeley, 2015; M. Hamilton & Rajaram, 2001b; Nelson & Schreiber, 1992; Paivio, 1967; Paivio et al., 1994). Accordingly, Wade and Vickery (2017) proposed that labels pertaining to an imageable person or entity (i.e., 'Self', 'Frog' and 'Snake') will create stronger and more memorable associations than less concrete labels (i.e., 'Stranger' or 'Other'). To test this hypothesis, they conducted a further experiment in which participants associated three shapes with the labels 'Self', 'Other', and 'Greg'. The label 'Greg' referred specifically to the experimenter that the participants had interacted with and was thus considered highly concrete. The results revealed similar processing benefits for the label 'Greg' as the label 'Self', supporting the idea that label concreteness, rather than threat, is driving the effects that arise in perceptual matching tasks. Thus, if perceptual matching is to be a valid test of social salience, caution must be taken to closely match selected labels across dimensions that might also influence salience. It remains unclear how motivating factors such as reward and emotional valence effect the salience of social stimuli.

5.3 The Present Study

In the present study I tested whether the emotional valence or 'likeability' of social stimuli biases prioritisation during perceptual processing. Using an adapted version of Sui et al's (2012) perceptual matching paradigm, participants were asked to associate three geometric shapes with: a) the self, b) an individual for whom they hold a favourable opinion, and c) an individual for which they hold an unfavourable opinion. Participants were then presented with a series of shape-label pairings and had to indicate if the pairing was correct or incorrect with respect to the previously learned associations.

In order to create appropriate stimuli, I selected labels that represent individuals for whom people often hold polarising opinions (positive or negative). To do so, I took inspiration from popular culture (Experiment 3) and politics (Experiment 4). In Experiment 3 participants associated three geometric shapes with the self, a hero from popular culture, and the respective villain. Because the participant pool consisted of the 2016 intake of undergraduate students, with a mean age of approximately 18 years, it was expected that most of that cohort would be familiar with the Harry Potter franchise. Therefore, in this experiment, the 'liked' character was the hero Harry Potter and the 'disliked' character was the villain Voldemort.

In Experiment 4, participants associated three shapes with the self, and two political adversaries for whom people tend to hold strong preferences for one over the other: the former President of the United States Barack Obama, and the incumbent President of the United States (at time of publication) Donald Trump. It is important to note that no assumptions were made as to the participants' preferences. Instead, at the end of Experiment 4 participants were asked to provide personal favourability ratings for each President. These ratings were then used to classify the stimuli as 'most-liked' and 'least-liked' for each participant post-hoc.

In line with previous perceptual matching studies, participants were asked to judge whether a series of shape-label pairings were correct or incorrect with respect to previously learned associations. The labels for Experiment 3 were 'You', 'Harry' and 'Voldemort', and the labels for Experiment 4 were 'You', 'Obama' and 'Trump'. As each of the labels referred to a recognisable individual they were considered to be highly concrete. Nevertheless, to control for imageability of associations and other label properties that might correlate with social salience, participants were also asked to judge shape-face pairings.

Face stimuli consisted of front-facing photos of the participants' own face, and similarly standardized images of the liked and disliked individuals. If prioritisation effects are found to be consistent across (text) label and (image) face pairings, we can assume that the shape stimuli activate global representations of the associated individual

(incorporating both their image and name or social label). Alternatively, if the results differ across stimulus modalities, then it may be that non-social properties of the words or images are the main drivers of prioritization effects.

Overall, the purpose of the present study was two-fold: i) to investigate whether the prioritisation of self-relevant information over other-relevant information depends on the likability of the 'other', and ii) to investigate the extent to which perceptual matching effects reflect the social rather than physical properties of the stimuli.

5.4 Experiment 3: Self vs. Hero vs. Villain associations

5.4.1 Method

5.4.1.1 Participants

Twenty-seven first year psychology undergraduate students (5 male, age: *M* = 18.6 years, range = 18–20 years) from the University of Warwick took part in the study. Data from two participants were excluded; one because of a data-recording error and one due to below chance performance (<50% correct responses across conditions). All participants confirmed that they were familiar with the Harry Potter franchise of books and movies. The sample size was decided based on previous perceptual matching studies investigating social salience (see Sui et al., 2012).

5.4.1.2 Stimuli and Apparatus

Displays were generated and responses recorded using custom programs running on a PC attached to a 19" LCD monitor running at a resolution of 1440 × 900 and 60Hz screen refresh rate. Stimuli were presented on a uniform grey background (RGB values = 128,128,128). On each trial one of three geometric shapes (triangle, square, or circle), presented in white (RGB values = 255,255,255) and subtending 4° × 4° at a viewing distance of 57 cm, was displayed 5° above a central fixation cross (1° ×

1°). In addition, either a text label or an image of a face was presented below the fixation cross (see Figure 8). Label stimuli consisted of the words "You" (2° × 1°), "Harry" (3° × 1°), and "Voldemort" (5.5° × 1°), presented in white in Arial font. Face stimuli consisted of a passport-style image of the participant (taken from their student ID), and similarly standardized face images of the characters Harry Potter and Voldemort from the Harry Potter movie franchise (face stimuli subtended 6.5° × 7°). All images were presented in greyscale and the edges were faded to the background colour in order prevent the image from having a distinctive shape. The distance between the centre of the shape, text label, or face image and the fixation cross was 5°.

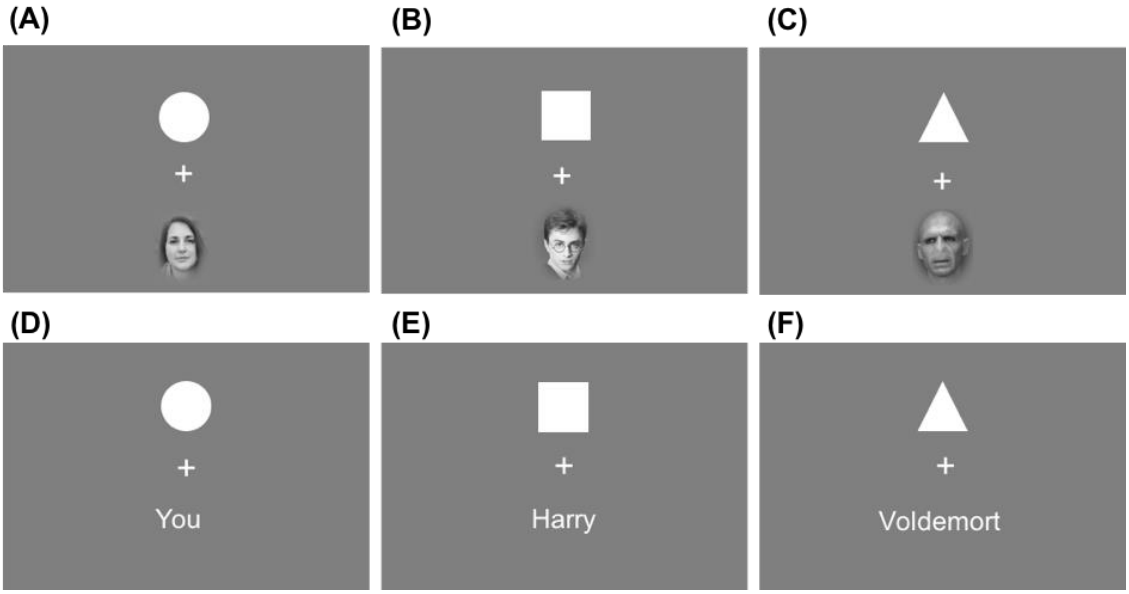

Figure 8. Example pairings between geometric shapes and the self, hero, and villain face stimuli (A-C) and self, hero, and villain text labels (D-E).

5.4.1.3 Design

The experiment used a within-participants 3 (shape category: self vs. hero vs. villain) × 2 (label modality: text vs. face) × 2 (shape-label match: matching vs. non-matching) design.

5.4.1.4 Procedure

There were two stages to the experiment. First, participants followed written instructions to encode three geometric shapes (triangle, square, and circle) as the self, the fictional hero Harry Potter, and his antagonist Voldemort. For example, a participant was told the following: "You are represented by a circle; Harry Potter is represented by a triangle; and Lord Voldemort is represented by a square". Shape-person associations were counterbalanced across participants. No shapes or faces were presented during this stage. Participants were given a total of 60s to memorise the associations.

The second stage involved a speeded perceptual matching task. Each trial began with the presentation of a central fixation cross for 500ms, after which a shape appeared above the fixation cross and either the image of a face or a text label appeared below it. Participants were required to press one of two keys (Z or M, counterbalanced across participants) to indicate whether or not the shape-label or shape-image pairing conformed to the previously learned associations. The stimuli remained on screen until the participant made a response or until a time-out of 4000ms had elapsed. At the end of each trial written feedback about the participant's response (correct, incorrect, or too slow) was presented at the centre of the screen for 500ms.

Participants completed a total of 720 trials over 6 blocks (120 trials per block). In each block self-, hero- and villain- associated shapes were presented equally often and in a random order. Each shape was paired equally often with a matched text label, a matched image, a mismatched text label, or a mismatched face image. For the mismatch trials, shapes were paired equally often with each incorrect face image or text label. Therefore, participants completed 60 trials per condition. Within each block half of the trials were matching and half non-matching, presented in random order.

5.4.2 Results

Table 3 presents mean correct RTs and accuracy data for Experiment 3. Responses faster than 200ms and missed responses (> 4000 ms) were excluded from analysis (<1% of trials). Accuracy was generally high across conditions (mean proportion correct = .89) and no speed-accuracy trade-offs were apparent.

Table 3. Mean Reaction Times and Accuracy as a Function of Match Condition (Matched vs. Unmatched), Shape Association (Self, Hero, and Villain) Social Label (Self, Hero, and Villain) and Label Modality (Face or Text) for Experiment 3.

Shape category	Label category	Correct Response	Label Modality	Mean RT (ms)	Accuracy
Self	Self	Match	Face	678.92 (68)	.93 (.08)
			Text	713.75 (80)	.90 (.10)
Hero	Hero	Match	Face	760.57 (101)	.90 (.08)
			Text	806.51 (88)	.87 (.08)
Villain	Villain	Match	Face	756.65 (99)	.88 (.12)
			Text	802.34 (101)	.85 (.10)
Self	Hero	Non-match	Face	833.77 (107)	.86 (.15).
			Text	866.71 (93)	.87 (.13)
Self	Villain	Non-match	Face	787.36 (86)	.92 (.09)
			Text	823.98 (100)	.92 (.10)
Hero	Self	Non-match	Face	790.31 (105)	.86 (.13)
			Text	831.41 (107)	.85 (.13)
Hero	Villain	Non-match	Face	839.37 (111)	.89 (.12)
			Text	894.88 (135)	.91 (.11)
Villain	Self	Non-match	Face	768.12 (92)	.93 (.10)
			Text	813.86 (105)	.90 (.10)
Villain	Hero	Non-match	Face	845.94 (115)	.90 (.13)
			Text	853.52 (123)	.87 (.12)

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses.

5.4.2.1 Accuracy

A signal detection approach (d') was used as a measure of response accuracy (see Macmillan & Creelman, 2004). For each shape association (Self, Hero, Villain) correct match responses were treated as 'hits', correct non-match responses were 'correct rejections', incorrect match responses were 'false alarms' and incorrect non-match responses were 'misses'.

A 3 (shape association: self vs. hero vs. villain) × 2 (label modality: text label vs. face image) within-subject ANOVA was conducted on d' scores. This showed that d' was significantly larger for shape-image pairings than shape-label pairings $F(1, 24) = 9.18$, $p = .006$, $\eta^2 = .28$, and there was also a main effect of shape association, $F(2, 48) = 4.05$, $p = .02$, $\eta^2 = .14$. However the two-way interaction did not approach significance, , $F(2, 48) = 1.23$, $p = .3$, $\eta^2 = .05$. Accordingly, pairwise analysis between shape associations was conducted with *d'* scores collapsed across label modality (Figure 9). d' was significantly larger for the self-shape than for the hero-shape, $t(24) = 2.81$, $p = .01$, $d = .56$,but all other pairwise comparisons were non-significant ($ts < 1.9$, $ps > .07$). Therefore, participants demonstrated greater accuracy when responding to the self-associated shape than the hero-associated shape, but there was no self-association advantage relative to the villain association.

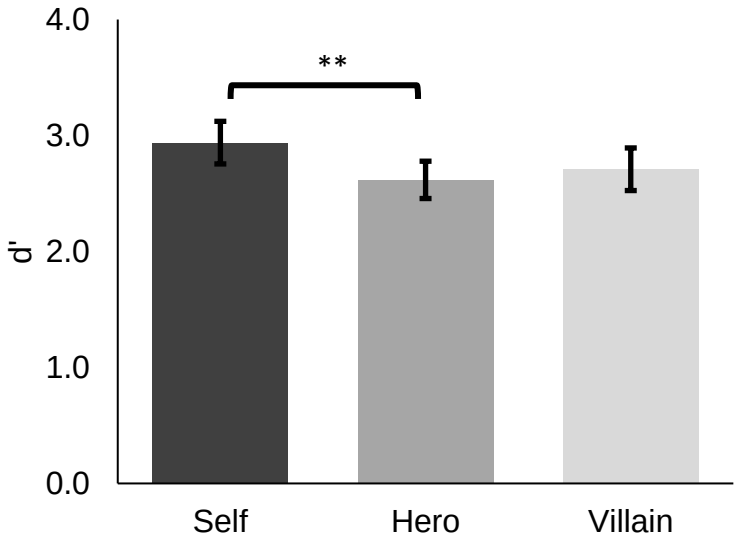

Figure 9. d' scores as a function of shape association. Error bars denote ± 1 standard error of the mean. Asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

5.4.2.2 RT data

An analysis of correct RTs was conducted using a 3 (shape association: self vs. hero vs. villain) × 2 (label modality: text label vs. face image) × 2 (label match: matching vs. non-matching) within-subjects ANOVA. This revealed a significant main effect of label match; participants were faster to verify a matching than a non-matching pair, $F(1, 24) = 105.13$, $p < .001$, $\eta^2 = .81$. Furthermore, there was a significant main effect of shape association, $F(2, 48) = 17.13$, $p < .001$, $\eta^2 = .42$, which was qualified by a significant interaction with -label match, $F(2, 48) = 16.93$, $p < .001$, $\eta^2 = .41$. Accordingly, RT data for the matching and non-matching trials were analysed separately below.

There was also a significant main effect of label modality; participants were faster to respond when the shape was paired with an image of a face than when it was paired with a text label, $F(1, 24) = 25.08$, $p < .001$, $\eta^2 = .51$. However, label modality did not interact with label match (matching or non-matching) or shape association, and the

three-way interaction was non-significant (*F*s< .17, *p*s > .5). As the pattern of results was consistent across shape-text and shape-face pairings, all further analysis was conducted on data collapsed across stimulus modality.

5.4.2.2.1 Matching trials.

Figure 10a shows mean correct RTs for trials in which the shape and label were correctly paired. A one-way ANOVA revealed a significant main effect of shape association, $F(2, 48) = 21.2$, $p < .001$, $\eta^2 = .47$, reflecting faster responses to the self-match pairs, than hero-match and villain-match pairs; $t(24) = 6.36$, $p < .001$, $d = 1.27$, and $t(24) = 4.82$, $p < .001$, $d = .96$, respectively. There was no significant difference in RTs between hero- and villain- shape-label associations, $t(24) = .35$, $p = .76$, $d = .06$.

5.4.2.2.2 Non-matching trials

Figure 10b shows the mean correct RTs for trials in which the shape and label were incorrectly paired for each of the three categories of shape association (self, hero, villain). A one-way ANOVA indicated no main effect of shape association, $F(2, 48) = 2.29$, $p = .11$ $\eta^2 = .09$.

5.4.2.2.3 Specific label effects.

On non-matching trials, a self-associated shape could form a non-matching pairing with either the hero label or with the villain label. Likewise, the villain associated shape could be presented with either self or hero labels, and the hero associated shape could be presented with either the self or villain labels. Thus, there were six possible non-matching pairings between shape association and label (self-villain, self-hero, hero-

self, hero-villain, villain-self and villain-hero). Therefore, for each shape (e.g., self, villain and hero associated shapes) I compared the RT difference between the two possible non-matching labels.

For the self-associated shape, RTs were significantly faster when it was paired with the villain label than when it was paired with the hero label, $t(24) = 4.96$, $p < .001$, $d = .99$. That is, people were quicker to determine that the self-associated shape did not match the villain label than the hero label. Participants were significantly faster to respond when the hero-shape had been paired with the self-label than the villain label, $t(24) = 2.41$, $p = .02$, $d = .48$, and when the villain-shape had been paired with the self-label than the hero-label, $t(24) = 4.47$, $p < .001$, $d = .89$.

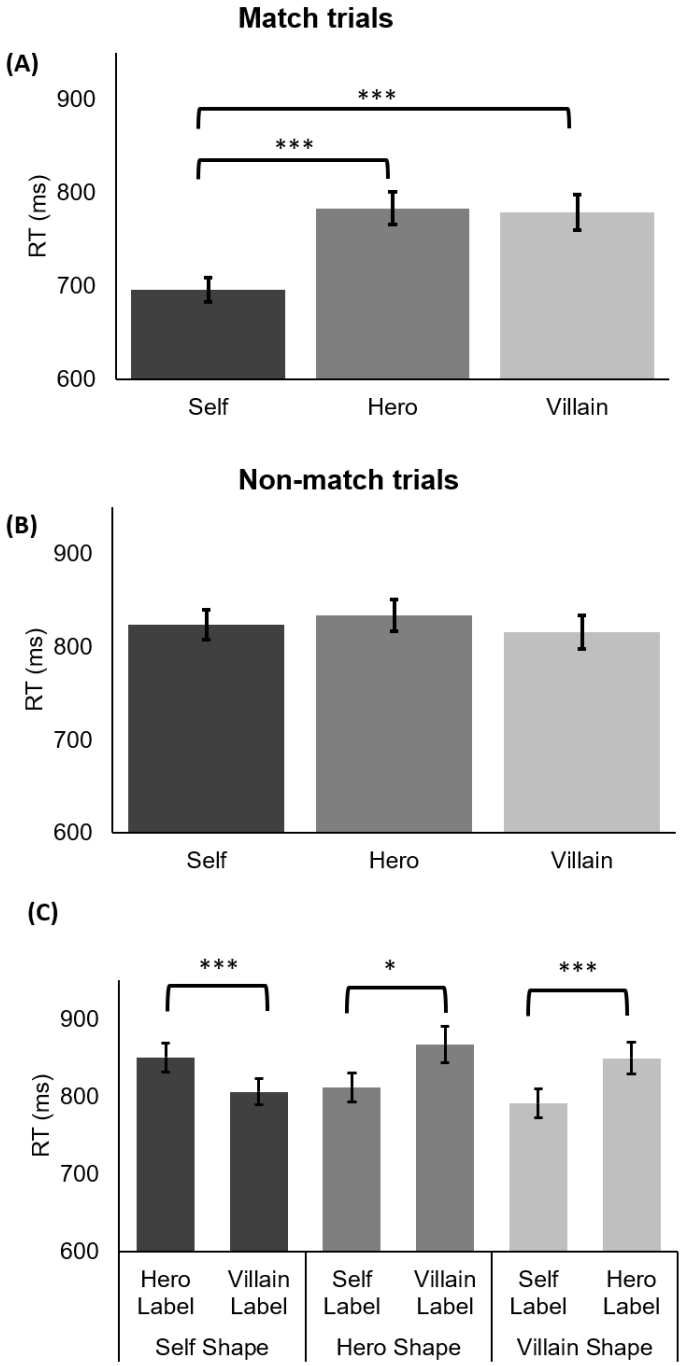

Figure 10. Mean correct RTs for matched (A) and nonmatching pairs (B) in Experiment 3, as a function of shape association (Self, Hero, Villain). Panel C shows mean correct RTs for each non-matching shape-label pairing. Error bars denote ± 1 standard error of the mean. Asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

5.4.3 Discussion

The results of Experiment 3 replicate the previously reported finding that newly formed associations between the self and an arbitrary stimulus can significantly benefit performance in a perceptual matching task (Sui et al., 2012). However, of most interest, no difference was found in the efficiency with which participants matched information to a hero or villain, suggesting that salience effects were determined by self-relevance rather than by positive or negative social evaluations.

Furthermore, prioritisation effects were consistent across label modality (i.e., whether the shape was paired with a text label or a face image). This finding suggests that the shape stimuli (square, triangle, circle) become assimilated within a *global* representation of an individual, rather than becoming tagged to a specific word or image. The fact that the same pattern of results was found for shape-face pairings as shape-text pairings suggests that perceptual matching effects are not driven solely by non-social properties of text-label stimuli (see Wade & Vickery, 2017).

Response times to non-matching pairings did not vary across shape categories. However, the non-match RTs for each shape association were influenced by which label the shape was paired with. For the hero and villain associated shapes, participants were significantly faster to verify a non-match if the label referred to the self. This is unsurprising given that self-referencing labels are likely to receive a similar prioritisation benefit as a self-associated shape.

Of particular interest, participants were faster to verify a non-match if the self-associated shape had been paired with a villain-label than if it had been paired with a hero label. While this may reflect the relative salience of each label, no significant difference was found between hero-match and villain-match responses, suggesting that one is not more salient than the other.

Alternatively, research in categorical perception has found that people are better able to distinguish between physically different stimuli that come from distinct categories (i.e. self vs other or liked vs. disliked) than physically different stimuli that are represented within the same category (Goldstone & Hendrickson, 2010a; Goldstone,

Lippa, & Shiffrin, 2001a). Following this logic, if stimuli are categorised as either 'self' and 'other', hero-villain pairings will appear 'less incorrect' than hero-self or villain-self pairings. Similarly, if stimuli are categorised as 'liked' or 'disliked' it is possible that participants perceive more incongruence between self-villain pairings than self-hero pairings. This categorisation aligns with research showing that perception is skewed to bolster positive self-regard (e.g., Heine, Lehman, Markus, & Kitayama, 1999). As such, I propose that there is greater representational overlap between the self and the 'liked other' than the self and the 'disliked other', resulting in the observed shape-label congruence effects.

In Experiment 4 this proposal was tested using stimuli representing real-world individuals. Participants were asked to associate three shapes with: i) the self, ii) Barack Obama and, iii) Donald Trump. After completing a perceptual matching task, participants provided personal favourability ratings for each president. These ratings were used to determine the categories of 'most liked' and 'least liked'.

5.5 Experiment 4: Self vs. Obama vs. Trump

5.5.1 Method

5.5.1.1 Participants

Twenty-seven first year psychology undergraduate students (3 male, age: M = 19.28 years, range = 18–26 years) from the University of Warwick took part in the study.

5.5.1.2 Stimuli and Procedure

The general procedure was the same as that of Experiment 3. However, participants were asked to associate three shapes (circle, triangle, and square) with the self, Barack Obama, and Donald Trump. Text-label stimuli consisted of the words

'You', 'Obama' and 'Trump'. Face stimuli consisted of a passport-style image of the participant (taken from their student ID), and similarly standardized face images of Barack Obama and Donald Trump (see Figure 11 for example shape-label and shape-face pairings).

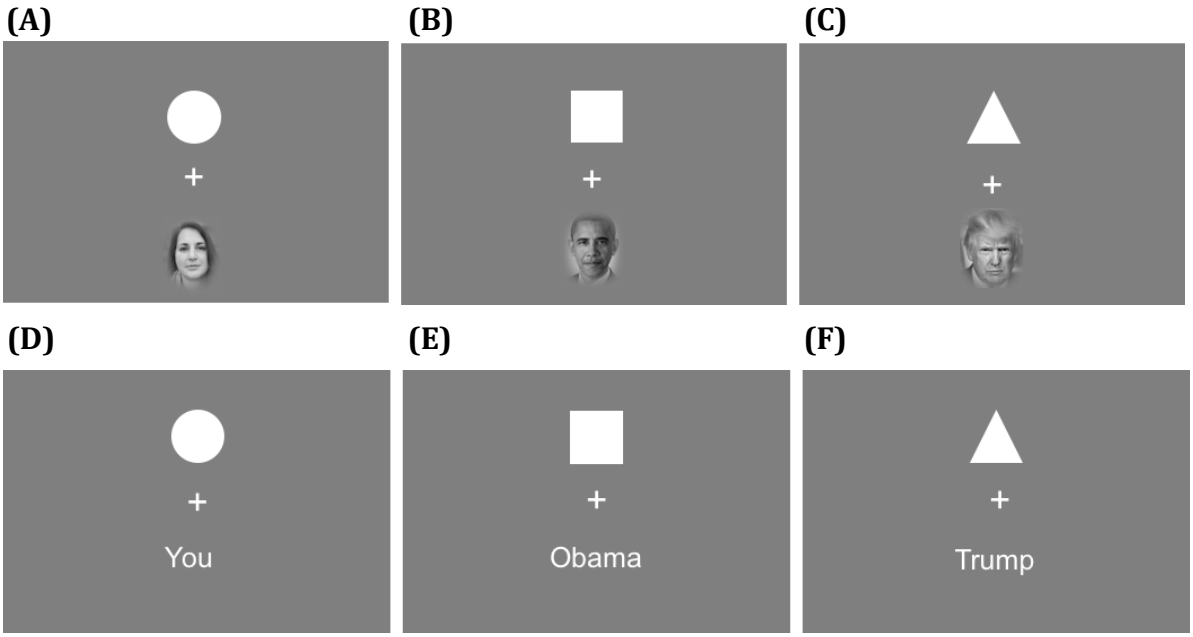

Figure 11. Example pairings between geometric shapes and the self, Obama, and Trump face stimuli (A-C) and self, Obama, and Trump text labels (D-E).

5.5.1.3 Favourability ratings

After performing the perceptual matching task, participants provided personal favourability ratings for each president by answering the following questions "How favourable or unfavourable is your view of Barack Obama?"/ "How favourable or unfavourable is your view of Donald Trump?". Responses were made on 5-point Likert scales ranging from -2 = very unfavourable, through -1 = somewhat unfavourable, 0 = neutral, +1 = somewhat favourable, to +2 = very favourable. Table 4 presents the

frequency of favourability scores for Trump, crossed with the favourability scores for Obama.

All participants rated Obama as more favourable than Trump, with the exception of one participant who gave both presidents a neutral rating. Obama's mean favourability score was +1.44 (SD = .7), whereas the mean favourability score for Trump was -1.7 (SD = .61), with a mean difference of 3.15 points (SD = 1.13). Due to the overall homogeneity of preferences, within the context of this experiment Donald Trump will be considered the least favourable of the presidents, and Barack Obama the most favourable.

Table 4. Cross tabulation of participants' favourability scores for Donald Trump and Barack Obama. Numbers represent frequency.

		Barack Obama				
		Very unfavourable	Somewhat unfavourable	Neutral	Somewhat favourable	Very favourable
Donald Trump	Very unfavourable	0	0	1	6	14
	Somewhat unfavourable	0	0	1	2	1
	Neutral	0	0	1	1	0
	Somewhat favourable	0	0	0	0	0
	Very favourable	0	0	0	0	0

5.5.2 Results

Table 5 presents mean correct RTs and accuracy data for Experiment 4. Responses faster than 200ms and missed responses (> 4000 ms) were excluded from analysis (<1% of trials). Accuracy was generally high across conditions (mean proportion correct = .92) and no speed-accuracy trade-offs were apparent.

Table 5. Mean Reaction Times and Accuracy as a Function of Match Condition (Matched vs. Unmatched), Shape Association (Self, Hero, and Villain) Social Label (Self, Hero, and Villain) and Label Modality (Face or Text) for Experiment 4.

Shape association	Label	Correct Response	Label Modality	Mean RT (ms)	Accuracy
Self	Self	Match	Face	722.57 (130)	.94 (.01)
			Text	748.74 (137)	.93 (.01)
Obama	Obama	Match	Face	804.25 (155)	.91 (.01)
			Text	855.17 (168)	.86 (.02)
Trump	Trump	Match	Face	834.78 (157)	.87 (.02)
			Text	863.00 (172)	.86 (.02)
Self	Obama	Non-match	Face	914.57 (189)	.92 (.02).
			Text	932.96 (180)	.92 (.02)
Self	Trump	Non-match	Face	859.67 (142)	.92 (.02)
			Text	900.35 (202)	.92 (.02)
Obama	Self	Non-match	Face	822.19 (152)	.91 (.02)
			Text	884.64 (162	.91 (.02)
Obama	Trump	Non-match	Face	918.71 (229)	.92 (.02)
			Text	953.54 (219)	.91 (.02)
Trump	Self	Non-match	Face	784.73 (136)	.95 (.01)
			Text	854.85 (137)	.91 (.02)
Trump	Obama	Non-match	Face	879.54 (153)	.93 (.02)
			Text	958.08 (218)	.90 (.02)

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses.

5.5.2.1 Accuracy

Figure 12 presents *d'* scores for each shape association (Self, Obama, Trump). A 3 (shape association: self vs. Obama vs. Trump) × 2 (label modality: text label vs. face image) within-subject ANOVA was conducted on *d'* scores. There was a main effect of label modality, $F(1, 26) = 5.54$, $p = .026$; , $\eta^2 = .18$. d' was significantly larger for shape-image pairings than shape-label pairings. There was also a main effect of shape association, $F(2, 52) = 16.87$, $p < .001$, $\eta^2 = .39$, however, the two-way interaction was non significant, $F(2, 48) = 1.23$, $p = .3$, $\eta^2 = .04$. Accordingly, pairwise analysis between shape associations was conducted with *d'* scores collapsed across label modality. *d'* was larger for the self-shape than the Obama-shape or Trump-shape; $t(26) = 5.36$, $p < .001$, $d = 1.03$, and $t(26) = 4.08$, $p < .001$, $d = 79$, respectively. No difference was found in *d'* scores for Obama and Trump associations, $t(26) = .46$, $p = .65$, $d = .09$. Therefore, responding was more accurate to self-associated shapes than to Obama or Trump associated shapes.

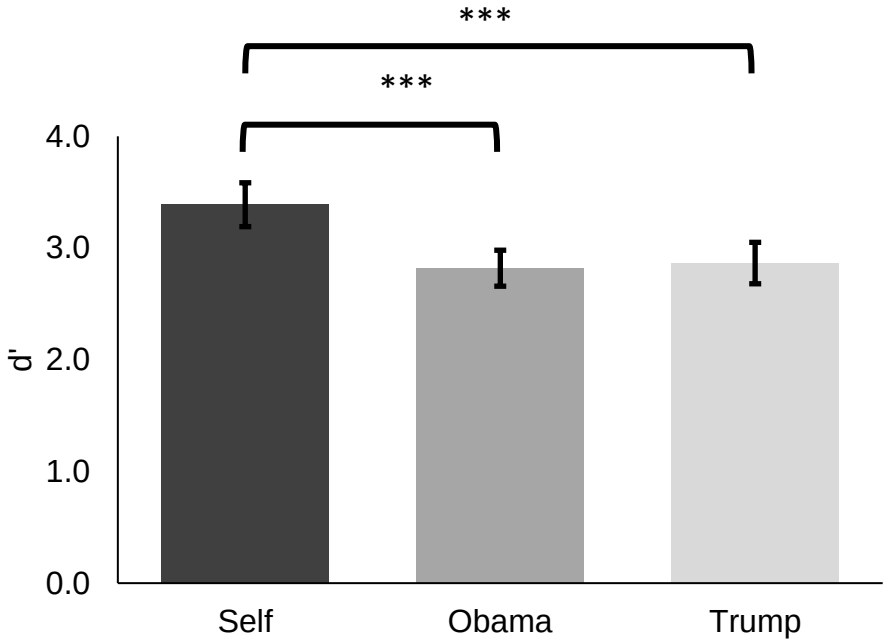

Figure 12. d' scores for Experiment 4. Error bars denote ± 1 standard error of the mean.

5.5.2.2 RT data

An analysis of correct RTs was conducted using a 3 (shape association: self vs. hero vs. villain) × 2 (label modality: text label vs. face image) × 2 (label match: matching vs. non-matching) within-subjects ANOVA. This revealed a significant main effect of label match; participants were faster to respond that a shape and label matched than mismatched, $F(1, 26) = 93.22$, $p < .001$, $\eta^2 = .78$. There was also a main effect of shape association, $F(2, 52) = 8.82$, $p < .001$, $\eta^2 = .25$, which was qualified by a significant interaction with label match, $F(2, 52) = 37.85$, $p < .001$, $\eta^2 = .59$. Henceforth, match and non- match trials will be analysed separately.

As with Experiment 3, there was a significant main effect of label modality on RTs, $F(1, 26) = 35.6$, $p < .001$, $\eta^2 = .58$. Participants were faster to respond when the shape was paired with a face than when it was paired with a text label. However, no interaction was found between label modality and shape association or shape-label

match, and there was no three-way interaction (*F*s< 1.37, *p*s > .5). As the pattern of results was consistent across shape-text and shape-face pairings, all further analysis was conducted on data collapsed across stimulus modality.

5.5.2.2.1 Match trials.

Figure 13a shows mean correct RTs for trials in which the shape and label were correctly paired. A one-way ANOVA revealed a significant main effect of shape association, $F(2, 52) = 28.61$, $p < .001$, $\eta^2 = .52$. Participants were significantly faster to respond to a matching shape-label pairing if the shape was associated with the self than if the shape was associated with either Obama or Trump; $t(26) = 5.34$, $p < .001$, $d = 1.03$ and $t(26) = 7.32$, $p < .001$, $d = 1.41$, respectively. No significant difference was found between Obama- and Trump- associations, $t(26) = 1.45$, $p = .16$, $d = .28$.

5.5.2.2.2 Non-matching trials.

Figure 13b shows the mean correct overall RTs for trials in which the shape and label were non-matching, for each of the three categories of shape association (self, Obama, Trump). ANOVA revealed a significant main effect of shape association, $F(2, 52) = 4.62$, $p = .01$, $\eta^2 = .15$; participants were faster to respond to a non-matching shape-label pairing when the shape was associated with Trump than when the shape was associated with either Obama or the self; $t(26) = 2.25$, $p = .03$, $d = .43$, and $t(26) = 3.26$, $p = .003$, $d = .43$, respectively. No reliable difference was found between Obama- and Self- associations, $t(26) = .68$, $p = .5$, $d = .13$.

5.5.2.2.3 Specific Label effects

*Figure 13*c presents the mean RTs for each shape association as a function of label type. On trials where the presented shape was associated with the self, RTs were shorter if the non-matching label was 'Trump' than 'Obama', $t(26) = 2.64$, $p = .01$, $d = .51$. When the presented shape was associated with Obama, RTs were faster if the label referred to the self than to Trump, $t(26) = 3.4$, $p = .002$, $d = .66$, and when the shape was associated with Trump responses were faster if the label referred to the self than Obama, $t(26) = 4.92$, $p < .001$, $d = .95$.

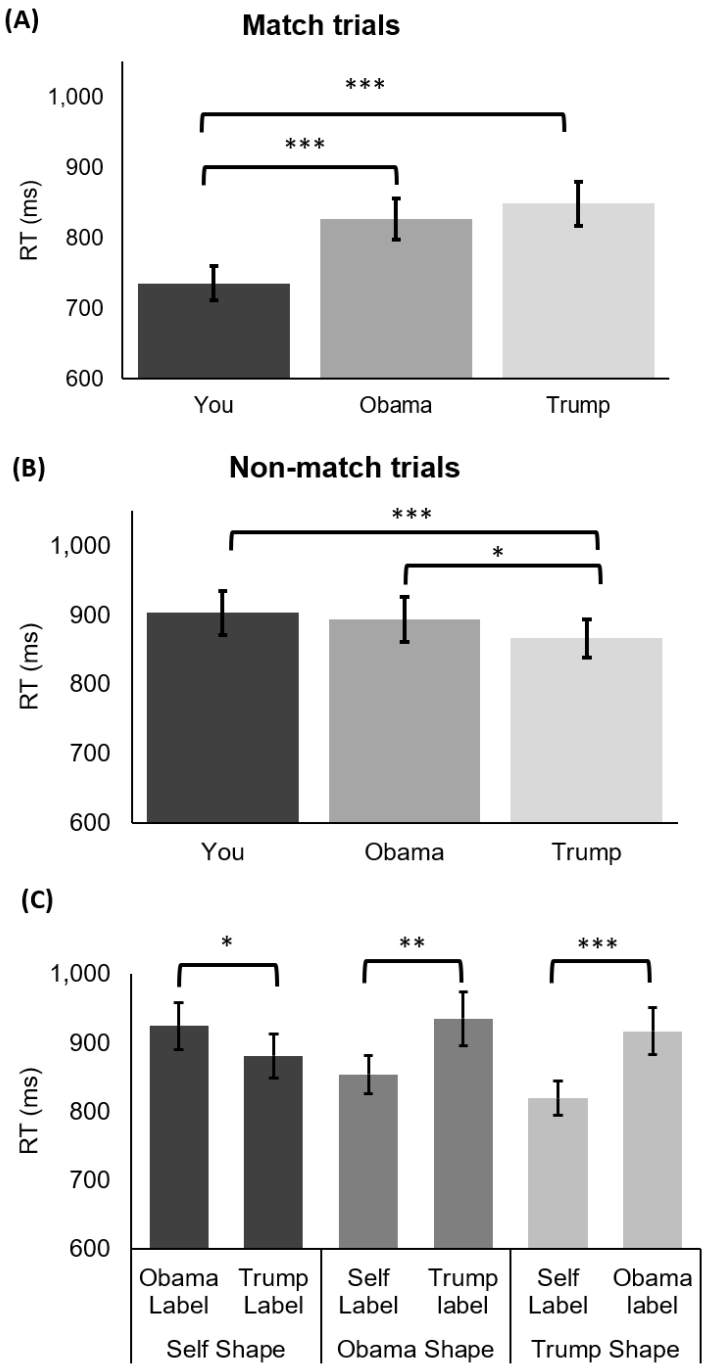

Figure 13. Mean correct RTs for matched (A) and non-matching pairs (B) in Experiment 4, as a function of shape association (Self, Obama, Trump). Panel C shows mean correct RTs for each non-matching shape-label pairing. Error bars denote ± 1 standard error of the mean. Asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

5.5.3 Discussion

The findings were largely consistent with those of Experiment 3. The same pattern of results was found across label modalities; it did not matter whether shapes were paired with a text label referring to one of the three individuals, or an image of their face. This is further confirmation that perceptual matching effects are not driven by physical properties of the labels, but rather by *who* the shape is associated with. I also found a reliable self-prioritisation effect; participants were faster to match stimuli to the self than either to Trump (the disliked other) or Obama (the liked other). Of note, no difference was found in match responses for Trump- and Obama- associated stimuli. This is further evidence that social stimuli are not prioritised for perceptual processing on the basis of 'likeability'.

As with Experiment 3, the label stimuli played a significant role in influencing non-match RTs. Participants were faster to verify a non-matched pairing when the Obama-shape was paired with a self-label, relative to when it was paired with a Trump-label. Likewise, participants were faster to verify a non-matched pairing when the Trump-shape was paired with a self-label, relative to Obama. Of note, when the Trump-shape was paired with the self-associated label (mismatched pair) participants were just as fast to respond as when the Trump-shape was paired with the Trump-label (matched pair). Consistent with Experiment 3, if the self-shape stimulus was paired with a label referring to the disliked other (Trump) responses were faster than if paired with a label referring to the liked other (Obama).

I argue that it is unlikely that label effects were simply a product of the relative salience of the label stimuli, otherwise it would have been expected that Trump-matches would be faster than Obama-matches. Alternatively, and in line with theories of categorical perception, I propose that the greater the perceived categorical overlap between the shape and label stimuli, the slower participants will be to separate them and verify a non-match (e.g., Heine, Lehman, Markus, & Kitayama, 1999).

Following this logic, we can assume that participants in our study perceived greater representational overlap between themselves and Obama than themselves and Trump. Given that the participants of Experiment 3 were also slower to distinguish between stimuli associated with the self and the liked-other (Harry) compared to the self and the disliked other (Voldemort), it is probable that 'likability' is the dimension on which the self and Obama overlap. This proposal is supported by research showing that people tend to have a positively-skewed self-representation (Heine et al., 1999; Ma & Han, 2010; L.A. Watson, Dritschel, Obonsawin, & Jentzsch, 2007).

5.6 General Discussion

The aim of this study was to determine the relative salience of neutral stimuli (geometric shapes) that had become associated with either the self, a liked other, or a disliked other. Four key findings were consistent across the two experiments. First, a reliable self-prioritisation effect was found. Participants were faster and more accurate when matching stimuli to the self than to either the liked or the disliked other. This result adds to an ever-growing body of evidence that self-biases are ubiquitous in cognition, relatively resistant to manipulation, and can be found across a range of contexts and cognitive domains (e.g., Keyes & Dlugokencka, 2014; Ma & Han, 2010; Stolte et al., 2017; Sui et al., 2012; Sui, Sun, Peng, & Humphreys, 2014; Sui et al., 2015; Sui & Humphreys, 2015b).

Second, we found no difference in the speed or accuracy with which participants matched shapes to the liked or disliked others. Thus, there was no evidence that the stimuli were prioritised for perceptual processing on the basis of likeability alone. Nonetheless, the self-prioritisation effect remained, which further supports the idea that self-biases are not fully equivalent to a positivity or reward bias (de Greck et al., 2008; Enzi et al., 2009; Ma & Han, 2010; Northoff & Hayes, 2011; Stolte et al., 2017; Sui & Humphreys, 2015d; Sui, Yankouskaya, et al., 2015).

Consistent with our findings, it is possible that the strength of the emotional response towards a stimulus matters more than the valence (A. K. Anderson, 2005; Dolcos, LaBar, & Cabeza, 2004). If the liked and disliked others evoke an equivalent magnitude of response in the observer, the associated stimuli would likely also have equivalent social salience. This could explain why some studies have found a processing bias for social stimuli conveying negative emotion or threat (Fox et al., 2000; Lang, Davis, & Öhman, 2000; Öhman et al., 2001; Tipples, Atkinson, & Young, 2002), whereas other studies have demonstrated a processing advantage for social stimuli conveying positive emotions (Kirita & Endo, 1995; Leppänen, Tenhunen, & Hietanen, 2003; Stolte et al., 2017) .

The third key finding was that when a shape was paired with the wrong label, non-matching trial RTs significantly differed depending on who the non-matching label referred to. On trials in which the shape was associated with one of the others (liked or disliked), non-match responses were fastest if the label referred to the self. On trials in which the shape was associated with the self, non-match responses were significantly faster if the label referred to the disliked-other than if the label referred to the liked-other. This pattern of results cannot be accounted for by the additive effects of shape and label salience alone, as the liked-other and disliked-other stimuli were found to be equally salient during match trials.

Instead, I propose that these findings align with theories of categorical perception. It has been well-established that perception is influenced by the way in which individuals categorically represent information (Goldstone & Hendrickson, 2010b). Once the physical differences have been accounted for, the ability to discriminate between perceptual items is reduced when items fall within rather than across category boundaries (Beale & Keil, 1995; Etcoff & Magee, 1992; Goldstone et al., 2001a). The category boundaries pertinent to the current research question were between 'self' and 'other', and 'liked' and 'disliked'.

There is reason to believe that the self-category has more representational overlap with the 'liked' category than the 'disliked' category. A diverse range of studies

have found that people have an inherent like for information associated with the self and a drive to enhance positive self-regard (Heine et al., 1999; Ma & Han, 2010; Markus & Kitayama, 1991; Northoff & Hayes, 2011). In line with this, the data from the current experiments show that participants found it harder to discriminate between two liked individuals (e.g., self and liked other) than between a liked other and a disliked individual (e.g., self and disliked other). Our data also suggest that it is harder to discriminate between two 'others' (e.g., liked-other and disliked-other) than to discriminate between the self and an 'other' (e.g. self and liked-other or self and disliked-other).

I propose that the perceptual matching task acted as a form of implicit bias test, as it measured the automatic associations a participant formed between the self, the disliked other, and the liked other. Traditional forms of implicit bias tests, such as The Implicit Association Test (IAT;Greenwald, McGhee, & Schwartz, 1998), have been used to determine implicit attitudes towards the self and others. For example, Greenwald et al. (1998) found that participants were faster to categorise themselves as positive rather than negative, consistent with the present findings. Ma and Han (2010) found that an implicit positive attitude towards one's own face mediated the self-advantage typically found in face recognition tasks. Furthermore, when the participants were asked to write down their personal flaws in order to prime self-concept threat, their implicit positive self-attitude was weakened, resulting in the amelioration of the self-face advantage. In the present study, the likeability of the self was not manipulated. However, it would be interesting to see whether the relative ease at which participants disassociated themselves from the disliked other would be reduced if this self-concept threat manipulation was introduced.

The final finding was that the pattern of results was consistent across shape-text and shape-label pairings. This is an indication that prioritisation effects were not driven by non-social properties of the specific labels, but that the shape stimuli had become assimilated within a global representation of an individual. Prioritisation was therefore determined by *who*, in a global sense, a stimulus was associated with. This consistency also reduces concerns that certain properties of text labels (such as word length or frequency) drive the selective prioritization effects (Sui et al., 2012).

Previous research has indicated that label concreteness can confound social salience effects (Wade & Vickery, 2017). Concreteness refers to the extent to which a word or concept relates to a perceptible entity, and has been associated with better memory recall and the formation of more robust associations (Campoy et al., 2015; M. Hamilton & Rajaram, 2001b; Nelson & Schreiber, 1992; Paivio et al., 1994). However, the stimuli in the present study were associated with well-known and imageable individuals (to the extent we could produce recognisable images of their faces), leading us to believe that all associations were highly concrete.

Overall, the results of the current study show that novel stimuli associated with the self are subject to processing benefits that are unlikely to reflect a general likeability bias. We have also demonstrated that representational overlap (here based on the likability dimension) between the self and others can influence the ease by which the self is distinguished from others. Our data supports the idea that the self is represented in a way that is distinct yet relatable to others (Decety & Sommerville, 2003b) but that the self-bias is not ultimately driven by likability alone.

In Chapter 6 I aimed to extend and replicate the findings from the above studies by asking participants to pick their own favourable and unfavourable others. Specifically, they had to select a current politician from whom they hold a favourable opinion of (Experiment 5) and unfavourable opinion of (Experiment 6).

6 Self vs. Stranger vs. Politician: Subjective favourability and Self-prioritisation

6.1 Abstract

In Chapter 5 I presented two experiments in which participants responded to stimuli associated with the self, a liked other, and a disliked other. In the present chapter I extended this work by asking participants to respond to novel stimuli associated with: (i) the self, (ii) an unspecified stranger, and (iii) a chosen politician. In Experiment 5 participants personally selected a politician for whom they held a favourable opinion. In Experiment 6 participants selected a politician for whom they held an unfavourable opinion. Having people choose their own stimuli in this way ensured that participants held truly opposing or dichotomous views of the social stimuli. A robust self-prioritisation effect was found in both experiments. Furthermore, participants were faster to match shapes to the politician than to the stranger. Cross-experiment comparisons revealed that the favourability of the politician did not influence the efficiency of responding to politician-associated stimuli. Accordingly, the favourability of the politician did not significantly influence the magnitude of the self-prioritisation effect.

6.2 Background

In Chapter 5, experiments demonstrated a robust prioritisation effect of self-associated stimuli over stimuli associated with a liked and a disliked other. Importantly,

no significant difference was found between the speed at which stimuli associated with a 'good' or 'liked' individual (e.g., Harry Potter) are perceptually processed compared to the speed at which stimuli associated with 'disliked' or 'bad' individuals (e.g., Voldemort) are processed. This pattern of results casts doubt over whether the self-prioritisation effect was driven by an inherent positive emotional evaluation of self-associated stimuli (see Stolte et al., 2017).

The experiments presented in the current chapter also investigate how the 'likeability' of social stimuli influence perceptual processing using Sui et al's (2012) perceptual matching paradigm. However, there are two key differences:

(1) Rather than having the 'favourable' and 'unfavourable' individual assigned to them, participants had to self-select a politician they either hold a favourable or unfavourable opinion of. This ensured the valence of favourability specific to the individual.

(2) In each experiment participants associated three novel shape stimuli with the self, a neutral other (a stranger), and EITHER a favourable (Experiment 5) or unfavourable politician (Experiment 6). Therefore, the key within-subject dependent variable was the strength of the emotional association with an individual rather than the valence of the association (favourable or unfavourable). The neutral stranger acts as control to see whether holding an emotional evaluation of an individual influences the perceptual processing of associated stimuli (regardless of salience).

If the *strength* of the emotional valence of the social stimuli influences perceptual processing, it would be expected that stimuli associated with a liked or disliked individual would be prioritised over stimuli associated with a neutral stranger. If the *valence* of the emotional valuation of the social stimuli influences perceptual processing, then it would be expected that the magnitude of the self-prioritisation effect would

differ depending on whether the chosen politician was liked or disliked. Thus, cross-experiment comparisons can be used to distinguish between the effects of strength and valence of the emotional value of social stimuli.

6.3 Experiment 5: Self vs. Stranger vs. Favourable politician

6.3.1 Method

6.3.1.1 Participants

Eighteen participants (three men; 18 to 21 years of age, M = 19.1) were recruited from the University of Warwick subject pool. All participants reported normal or corrected-to-normal vision and received course credit for their participation.

6.3.1.2 Stimuli

Displays were generated and responses recorded using custom programs running on a PC attached to a 19" LCD monitor running at a resolution of 1440 × 900 and 60Hz screen refresh rate. All stimuli were presented in white (RGB values = 255,255,255) on a uniform grey background (RGB values = 128,128,128). On each trial one of three geometric shapes (triangle, square, or circle) each subtending 3.8° × 3.8° at a viewing distance of 57 cm, was presented above a central fixation cross (0.8° × 0.8°). In addition, a text label was presented below the fixation cross. Label stimuli consisted of the words "Self" (2.4° × 1°), "Politician" (4.5° × 1°), and "Stranger" (5.5° × 1°), presented in Arial font. The distance between the centre of the shape or the label stimuli and the fixation cross was 3.5°.

6.3.1.3 Procedure

Participants were asked to think of a specific politician for whom they held a favourable opinion. For instance, participants were asked "Please think of a specific politician for whom you hold a positive opinion. This can be a current or recent politician (active in the last 10 years)".

Participants were presented with written instructions to encode three geometric shapes (triangle, square, and circle) as the self, the chosen politician, and a hypothetical stranger. For example, a participant could be given the following instructions: "In this task you are represented by a triangle. The chosen politician is represented by a circle, and a stranger is represented by a square." Shape-person associations were counterbalanced across participants. Participants had 60 seconds to encode the shape-person associations prior to the start of the perceptual matching task. No visual stimuli were presented on the display at this stage.

Immediately following the encoding stage, participants began the perceptual matching task. Each trial began with the presentation of a central fixation cross for 500ms. Next, a random pairing of a shape stimulus (triangle, circle, or square) and a label stimulus ('Self', 'Politician' or 'Stranger') was presented on the display for 100ms. The pairing either conformed to the just-learned associations (i.e., a matching pair), or were a recombination of a shape with a non-matching label (i.e., a non-matching pair). A blank display was then presented for up to 2000ms, or until participants had responded. Participants responded by indicating whether the shape-label pair was matching or non-matching by pressing either Key Z or Key M on a QWERTY keyboard. The assignment of each key to 'matching' or 'non-matching' responses was counterbalanced across participants. After the response window ended, feedback consisting of 'correct response' or 'incorrect response' was presented at the centre of the display for 500ms.

Following 12 practice trials, each participant completed a total of 480 trials divided equally over four blocks. Within each block, the self-, politician- and stranger-associated shapes were paired equally often with a matching and non-matching label. The pairings were presented in a random order. Accordingly, there were 80 trials in each

condition (self-matched, self-nonmatching, politician-matched, politician-nonmatching, stranger-matched, and stranger-nonmatching).

6.3.2 Results

Table 6 presents mean correct RTs and accuracy data for Experiment 5. Responses faster than 200ms and missed responses (> 2000 ms) were excluded from analysis (<1% of trials across both experiments). Accuracy was generally high across conditions (mean proportion correct = .82). No speed-accuracy trade-off was apparent.

Table 6. Mean correct RTs and Accuracy as a Function of Shape Association (Self, Politician, and Stranger), Label (Self, Politician, and Stranger), and Response Type (Match, Mismatch) for Experiment 5 (favourable politician).

Shape Association	Label	Response Type	RT	Accuracy
Self	Self	Match	555.02 (60)	.91 (.07)
	Politician	Mismatch	712.07 (109)	.87 (.09)
	Stranger	Mismatch	731 (106)	.83 (.11)
Politician	Politician	Match	655.06 (76)	.86 (.07)
	Self	Mismatch	696.27 (111)	.83 (.10)
	Stranger	Mismatch	786.02 (122)	.82 (.10)
Stranger	Stranger	Match	720.23 (93)	.77 (.05)
	Self	Mismatch	683.85 (110)	.86 (.10)
	Politician	Mismatch	734.66 (108)	.83 (.11)

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses.

6.3.2.1 Accuracy

A signal detection approach (d') was used as a measure of response accuracy. For each shape association (Self, Politician, Stranger) correct match responses were treated as 'hits', and correct non-match responses were 'correct rejections'. A larger d' indicates greater sensitivity of responding. Figure 14 presents mean d' scores for each shape association for Experiment 5.

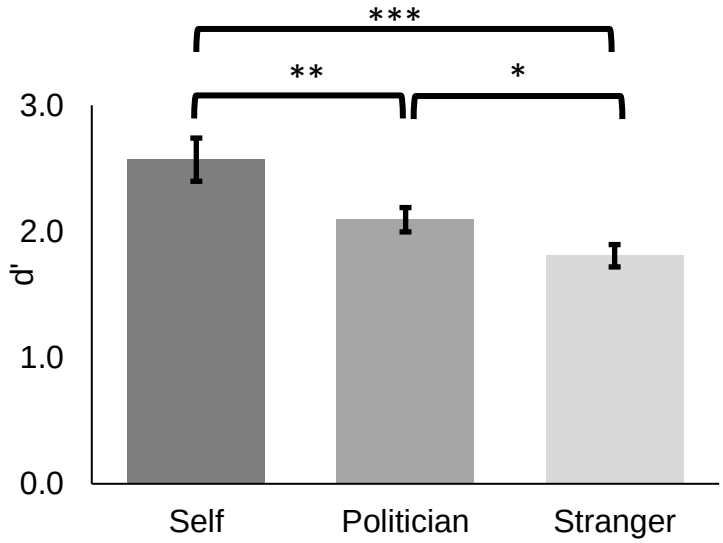

Figure 14. d' scores as a function of shape association for Experiment 5 Error bars denote ± 1 standard error. Asterisks denote a significance levels (* p ≤ 0.05, ** p ≤ 0.01, *** p ≤ 0.001).

There was a main effect of shape association on response sensitivity, $F(2, 34) = 15.42$, $p < .001$, $\eta^2 = .48$. Responses were more accurate (larger d') when the shape stimulus was associated with the self than with either the politician or the stranger; $t(17) = 3.28$, $p = .004$, $d = .77$ and $(17) = 4.94$, $p < .001$, $d = 1.17$, respectively. Also, participants were significantly more accurate when responding to the politician-associated shape than the stranger-associated shape, $t(17) = 2.53$, $p = .02$, $d = .59$.

6.3.2.2 RT data

A 3 (shape association: self, politician, stranger) × 2 (pairing type: matched, non-matched) within-subjects analysis was conducted on mean correct RTs. There was a main effect of pairing type, with responses to matched pairs being faster than responses to non-matched pairs, $F(1, 17) = 34.09$, $p < .001$, $\eta^2 = .67$. There was also a main effect of shape association, $F(2, 34) = 32.11$, $p < .001$, $\eta^2 = .65$, which was qualified by a

significant interaction with pairing type, $F(2, 34) = 41.61$, $p < .001$, $\eta^2 = .71$. Accordingly, match and non-match RTs were analysed separately below.

6.3.2.2.1 Match RTs

*Figure 15*a presents match RTs for Experiment 5. There was a main effect of shape association on match RTs, $F(2, 34) = 61.36$, $p < .001$, $\eta^2 = .78$. Pairwise analysis revealed that self-match RTs were significantly shorter than politician-match or stranger-match RTs; $t(17) = 6.56$, $p < .001$, $d = 1.54$ and $t(17) = 9.69$, $p < .001$, $d = 2.26$, respectively. Also, politician-match RTs were significantly shorter than stranger-match RTs, $t(17) = 5.25$, $p < .001$, $d = 1.24$.

6.3.2.2.2 Non-match RTs

*Figure 15*b presents non-match RTs for Experiment 5. There was no effect of shape association on non-match RTs, $F(2, 34) = 2.76$, $p = .08$, $\eta^2 = .14$.

6.3.2.2.3 Specific Label effects

On non-matching trials each shape category (self, politician, stranger) could be paired with one of two non-matching labels (see *Figure 15*c). Pairwise comparisons were conducted to determine whether the identity of the non-matching label influenced RTs for each shape category. When the presented shape was associated with the self, there was no significant difference in RTs between the politician label and the stranger label, $t(17) = 1.03$, $p = .32$, $d = .24$. When the presented shape was associated with the politician, RTs were significantly shorter when the label referred to the self than the stranger, $t(17) = 6.42$, $p < .001$, $d = 1.51$. When the presented shape was associated with

the stranger, RTs were significantly shorter when the label referred to the self than the politician, $t(17) = 2.58$, $p = .02$, $d = .61$.

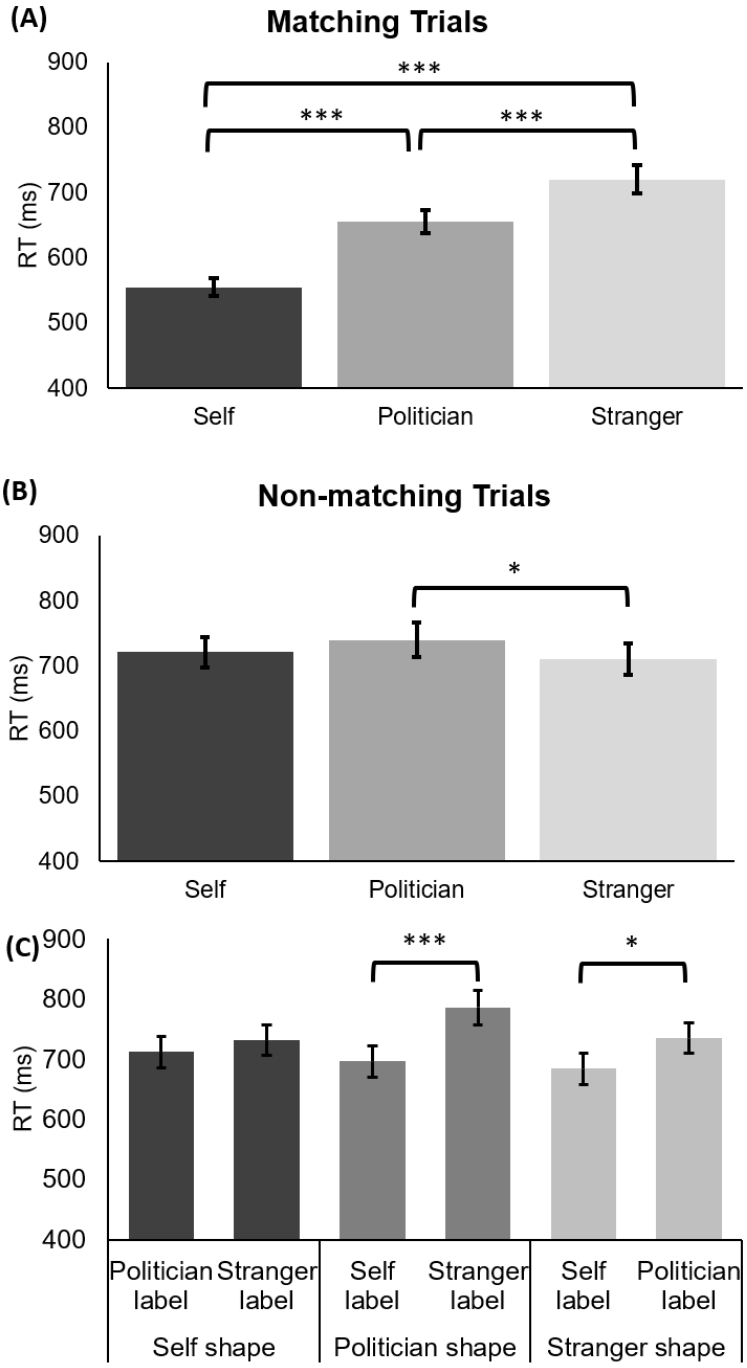

Figure 15. Mean correct RTs for matched (Panel A) and non-matching pairs (Panel B) in Experiment 5, as a function of shape association (Self, Politician, Stranger). Panel C shows mean correct RTs for each non-matching shape-label pairing. Error bars denote ± 1 standard error of the mean. Asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

6.3.3 Discussion

Participants were more efficient (shorter RTs and greater accuracy) at responding to a novel visual stimulus associated with the self than a novel visual stimulus associated with a stranger or a personally-selected favourable politician. In addition, participants were more efficient at responding to stimuli associated with the favourable politician than the stranger. The robust self-prioritisation effect found in Experiment 5 is consistent with that found in the experiments of Chapters four and five, as well as the vast majority of studies using Sui et al's (2012) perceptual matching paradigm to investigate the relationship between social salience and perceptual salience(Stolte et al., 2017; Sui et al., 2012; Sui & Humphreys, 2015b; Sui, Yankouskaya, et al., 2015).

When the presented shape-label pairings did not match, the social affiliation of the shape stimuli had no overall effect on the speed of responses. However, when the politician-associated shape was paired with a non-matching label, responses were faster if the label referred to the self than the stranger. Also, when the stranger-associated shape was paired with a non-matching label, responses were faster if the label referred to the self than to the politician. This pattern of results may reflect a general processing advantage for the self-associated shape over the other associations. However, it is also well-established that the way in which items are categorized influences perceptual processing, with between-category contrasts being easier to perform than within-category contrasts (Goldstone & Hendrickson, 2010a; Goldstone, Lippa, & Shiffrin, 2001b). Therefore, these label effects may reflect faster responding for self-other discriminations than other-other discriminations.

Overall, the finding that stimuli associated with a liked politician were processed more efficiently than stimuli associated with a stranger is an indication that social stimuli may be prioritized for perceptual processing based on the likeability of the referent. However, there are a number of other reasons why the politician shape would be more salient than the stranger shape:

(i) As participants personally selected the politician, the associated stimulus has more personal significance.

(ii) Asking participants to think of a politician generated deeper and more elaborative processing when encoding the shape-label associations (Craik & Tulving, 1975; Eysenck & Eysenck, 1979).

(iii) the instructions were particularly transparent in drawing attention to the politician category and emphasizing its importance.

(iv) Associations with a well-known specific politician will be more concrete than associations with a unspecified stranger.

To address the above issues, a comparison experiment (Experiment 6) was conducted in which participants thought of a politician for whom they held an *unfavourable* opinion. Therefore, by comparing the data from Experiment 5 and Experiment 6, it will be possible to determine whether the relative likability of the politicians influences the prioritisation of social stimuli for perceptual processing.

6.4 Experiment 6: Self vs. Stranger vs. Unfavourable politician

In Experiment 6 participants performed a perceptual matching task with stimuli associated with the self, a stranger and a personally-selected politician for whom they held an unfavorable opinion.

6.4.1 Method

6.4.1.1 Participants

Eighteen participants (five men; 18 to 21 years of age, M = 18.7) were recruited from the University of Warwick subject pool. All participants reported normal or corrected-to-normal vision and received course credit for their participation.

6.4.1.2 Stimuli and Procedure

The stimuli and procedure were identical to those of Experiment 5, with one key exception. At the beginning of the experiment participants were asked to specify a politician for whom they held an unfavorable opinion. One of the three shape stimuli would then be used to represent than specific politician.

6.4.2 Results

Table 7 presents mean correct RTs and accuracy data for Experiment 6. Responses faster than 200ms and missed responses (> 2000 ms) were excluded from analysis (<1% of trials across both experiments). Accuracy was generally high across conditions (mean proportion correct = .83). No speed-accuracy trade-off was apparent.

Table 7. Mean correct RTs and Accuracy as a Function of Shape Association (Self, Politician, and Stranger), Label (Self, Politician, and Stranger), and Response Type (Match, Mismatch) for Experiment 6 (unfavourable politician).

Shape Association	Label	Response Type	RT	Accuracy
Self	Self	Match	578.15 (82)	.88 (.11)
	Politician	Mismatch	720.04 (76)	.86 (.11)
	Stranger	Mismatch	778.17 (98)	.81 (.13)
Politician	Politician	Match	658.96 (106)	.84 (.12)
	Self	Mismatch	679.28 (70)	.83 (.18)
	Stranger	Mismatch	799.89 (142)	.78 (.17)
Stranger	Stranger	Match	704.08 (88)	.72 (.16)
	Self	Mismatch	719.20 (104)	.86 (.16)
	Politician	Mismatch	756.67 (93)	.80 (.18)

Note. RT = reaction time; Accuracy = proportion correct. Standard deviations appear within parentheses.

6.4.2.1 Accuracy

Figure 16 presents mean d' scores for each shape association for Experiment 6. There was a main effect of shape association on response sensitivity, $F(2, 34) = 12.06$, $p < .001$, $\eta^2 = .42$. Responses were significantly more accurate (larger d') when the shape stimulus was associated with the self than either the politician or the stranger; $t(17) = 2.79$, $p = .001$, $d = .64$ and $.t(17) = 5.23$, $p < .001$, $d = .123$, respectively. Also, participants were significantly more accurate when responding to the politician-associated shape than the stranger-associated shape, $t(17) = 2.17$, $p = .04$, $d = .51$.

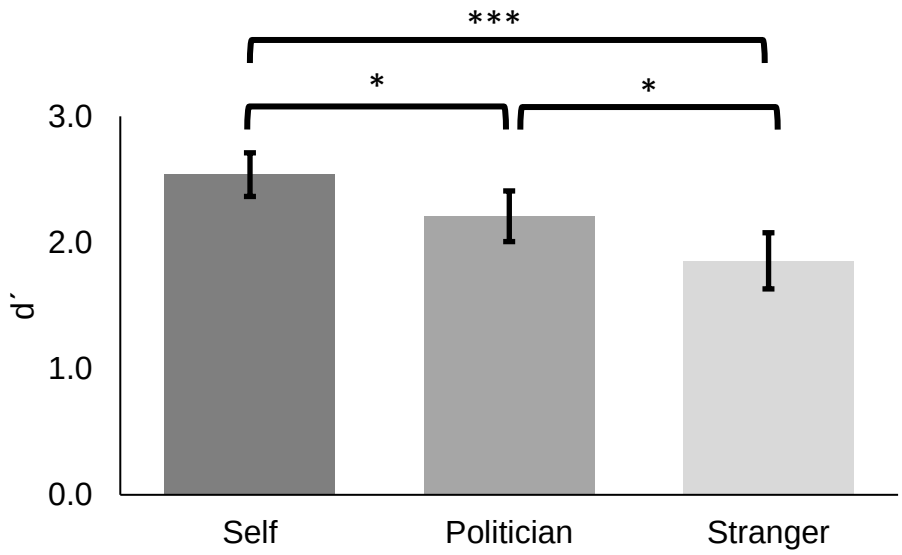

Figure 16. d' scores as a function of shape association for Experiment 6. Error bars denote ± 1 standard error. Asterisks denote a significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

6.4.2.2 RT data

A 3 (shape association: self, politician, stranger) × 2 (pairing type: matched, non-matched) within-subjects analysis was conducted on mean correct RTs. There was a main effect of pairing type, with match responses being faster than non-match responses, $F(1, 17) = 114.45$, $p < .001$, $\eta^2 = .87$. There was also a main effect of shape association, $F(2, 34) = 15.65$, $p < .001$, $\eta^2 = .48$, which was qualified by a significant interaction with pairing type, $F(2, 34) = 19.09$, $p < .001$, $\eta^2 = .53$. Accordingly, match and non-match RTs were analysed separately below.

6.4.2.2.1 Match RTs

*Figure 17*a presents the mean match RTs. There was a main effect of shape association on match RTs, $F(2, 34) = 24.86$, $p < .001$, $\eta^2 = .59$. Pairwise analysis revealed that self-match RTs were significantly shorter than politician-match or stranger-match RTs; $t(17) = 5.37$, $p < .001$, $d = 1.27$ and $t(17) = 6.41$, $p < .001$, $d = 1.51$, respectively. Also, politician-match RTs were significantly shorter than stranger-match RTs, $t(17) = 2.35$, $p = .03$, $d = .55$.

6.4.2.2.2 Non-match RTs

*Figure 17*b presents the mean non-match RTs. There was no main effect of shape association on non-match RTs, $F(2, 34) = 0.81$, $p = .45$, $\eta^2 = .05$.

6.4.2.2.3 Specific Label effects

*Figure 17*c presents mean non-match RTs for each shape-label combination. When the presented shape was associated with the self, non-match responses were significantly faster when the label referred to the politician than the stranger, $t(17) = 3.96$, $p = .001$, $d = .93$. When the presented shape was associated with the politician, non-match responses were significantly faster when the label referred to the self than the stranger, $t(17) = 4.17$, $p < .001$, $d = .98$. There was no significant label effect when the shape associated with the stranger, $t(17) = 1.84$, $p = .08$, $d = .44$.

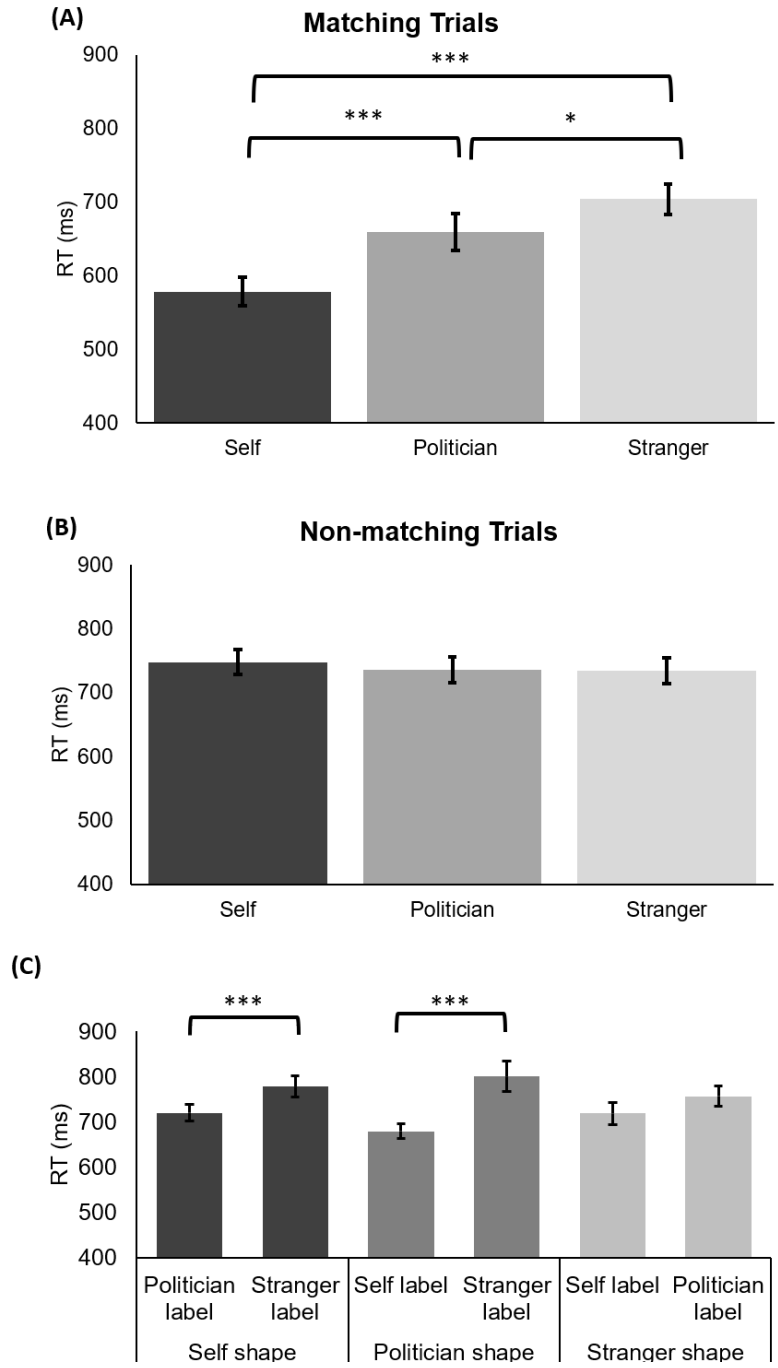

Figure 17. Mean correct RTs for matched (Panel A) and non-matching pairs (Panel B) in Experiment 6, as a function of shape association (Self, Politician, Stranger). Panel C shows mean correct RTs for each non-matching shape-label pairing. Error bars denote ± 1 standard error of the mean. Asterisks denote significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

6.4.3 Discussion

As in Experiment 5, a robust self-prioritisation effect was found. Participants were significantly faster and more accurate when matching shape stimuli to the self than to either the stranger or the unfavourable politician. Importantly, participants were also faster and more accurate when matching stimuli to the unfavourable politician than to the stranger.

When shape-label pairings did not match, there was no overall effect of shape-association (self, politician, stranger) on response times. However, there were specific label effects. When the presented shape was associated with the unfavourable politician, participants were faster to indicate that the shape and label pairing did not match if the label referred to the self than the stranger. Also, when the presented shape was associated with the self, participants were faster to indicate that the shape and label pairing did not match if the label referred to the unfavourable politician than the stranger. This pattern of results suggests that participants were faster to make self-politician discrimination judgements than self-stranger discrimination judgements. It is possible that the greater the categorical or representational overlap between two individuals, the harder it is to discriminate between associated stimuli. It is a well-established phenomenon that cognitive and perceptual processes become distorted in order to maintain positive self-regard (Blaine & Crocker, 1993; Forsyth, 2008). Therefore, it is likely that participants are driven to perceive themselves as having less in common with the unfavourable politician than the neutral stranger.

To further understand how the favourability of social stimuli influences perceptual processing, the data from Experiment 5 (self vs. stranger vs. favourable politician) was compared with that of the current experiment (self vs. stranger vs. unfavourbale politician).

6.5 Cross Experiment Comparisons

Cross-experiment comparisons were conducted to determine directly whether the favourability of the politician influenced the prioritisation of self, politician, and stranger associated stimuli. Analyses were conducted using mixed ANOVAs with shape association (self, partner, stranger) as a within-subjects variable and favourability of politician (favourable, unfavourable) as a between-subjects variable.

6.5.1 Accuracy

There was no main effect of politician favourability on d' scores, $F(1,34) = .05$, $p = .83$, $\eta^2 = .001$, and no significant interaction between shape association and politician favouribility, $F(2,68) = .27$, $p = .76$, $\eta^2 = .005$.

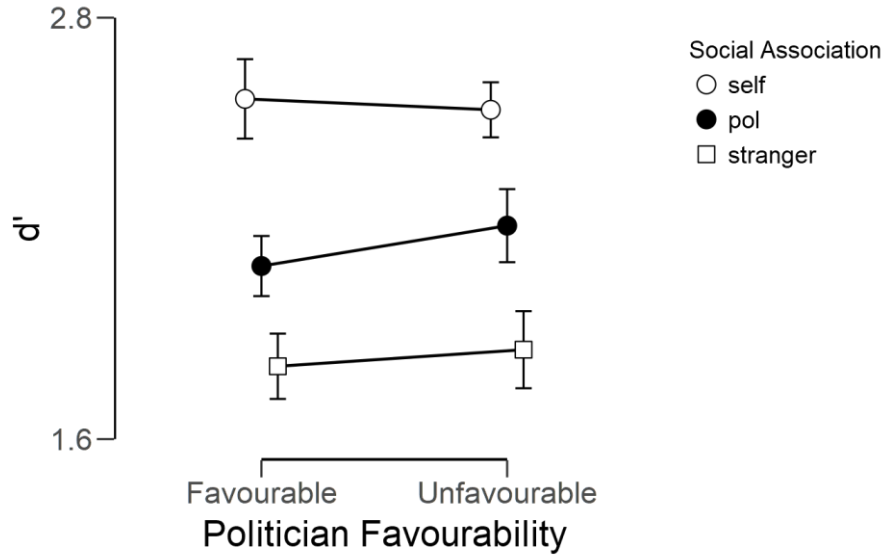

Figure 18. *d'* scores as a function of shape association across Experiment 5 (favourable politician) and Experiment 6 (unfavourable politician). Error bars denote ± 1 standard error. Asterisks denote a significance levels (* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$).

6.5.2 RT data

Figure 19 displays the mean match RTs for self-, politician-, and stranger-associated shapes as a function of politician favourability. *Figure 20* shows the equivalent non-match RTs. Match and non-match RTs will be analysed separately.

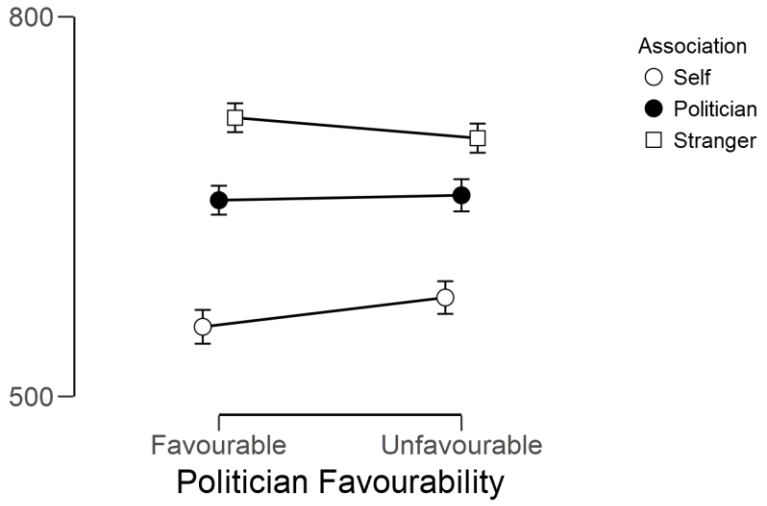

Figure 19. Mean correct RTs for match trials as a function of shape association (self, politician, stranger) when the politician was favourable (Experiment 5) and unfavourable (Experiment 6). Error bars denote ± 1 standard error.

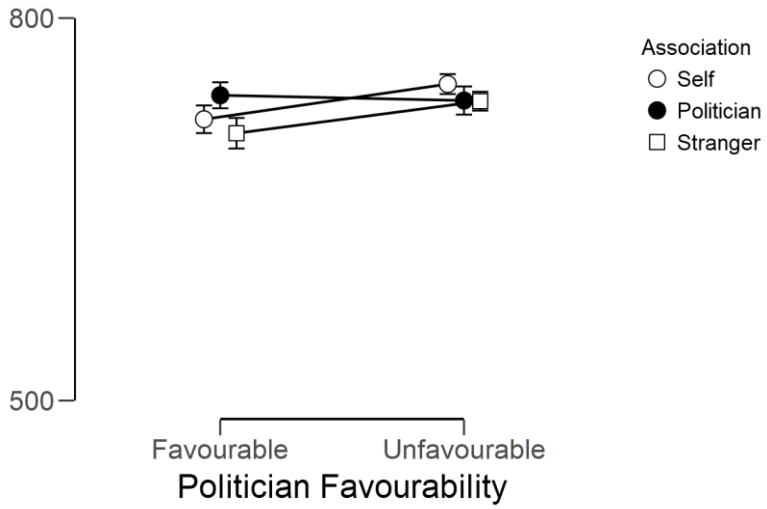

Figure 20. Mean correct RTs for non- match trials as a function of shape association (self, politician, stranger) when the politician was favourable (Experiment 5) and unfavourable (Experiment 6). Error bars denote ± 1 standard error.

6.5.2.1 Match RTs

There was no significant main effect of politician favourability on match RTs, $F(1,34) = .02$, $p = .87$, $\eta^2 = .001$. and no interaction was found between shape association and politician favourability, $F(2,68) = 1.4$, $p = .26$, $\eta^2 = .01$.

6.5.2.2 Non-match RTs

No main effects of politician favourability or shape association (self, politician, stranger) were found on non-match RTS; $F(1,34) = .29$, $p = .6$, $\eta^2 = .008$ and $F(2,68) = 1.68$, $p = .19$, $\eta^2 = .05$, respectively. There was no significant interaction between politician favourability and shape association, $F(2, 68) = 2.03$, $p = .14$, $\eta^2 = .05$.

6.5.2.3 Specific label effects on self-non-match RTs

Figure 21 shows the mean non-match RTs for the self-associated shape as a function of label type and partner presence. Previous analysis revealed that when the politician was unfavourable (Experiment 6), participants were significantly slower to make a non-match response if the self-associated shape was presented with a label referring to the partner than the stranger. However, this label effect was not found when the politician was favourable (Experiment 5). Cross-experiment analysis revealed that the interaction between politician favourability (favourable, unfavourable) and label type (politician, stranger) was nonsignificant, $F(1,34) = 2.79$, $p = .104$, $\eta^2 = .06$.

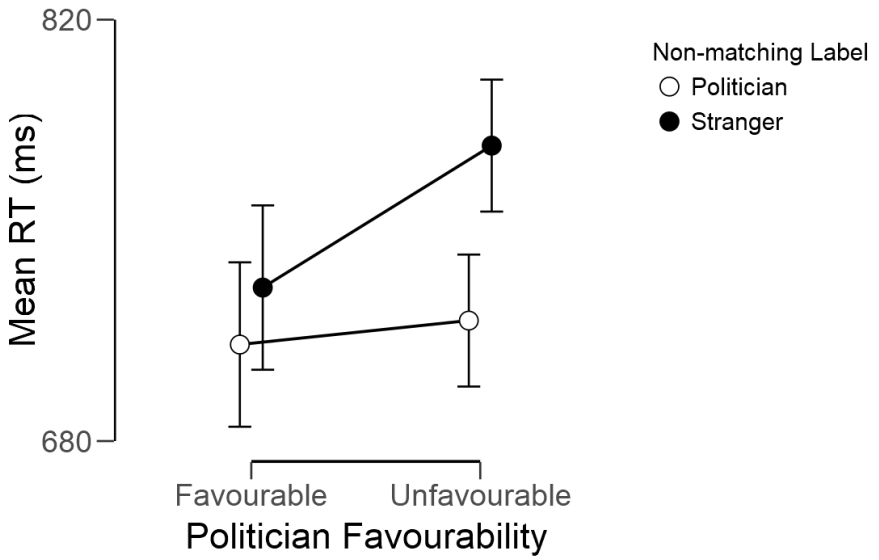

Figure 21. Mean non-match RTs for the self-associated shape as a function of label type and politician favourability. Error bars denote ± 1 standard error.

6.5.3 Discussion of cross-task comparisons

Overall, RTs and accuracy were not found to differ significantly between Experiment 5 (self vs. stranger vs. favourable politician) and Experiment 6 (self vs. stranger vs. unfavourable politician).

6.6 General Discussion

Previous studies that have attempted to disassociate self-bias from other forms of bias have found that a participant's sensitivity to the reward value (Sui & Humphreys, 2015d), or emotional valence (Stolte et al., 2017), of a visual stimulus does not predict their sensitivity to the self-relevance of a perceptual stimulus. This was taken to suggest that self-bias and positivity/reward biases have distinct neural origins. Rather than taking a correlational approach, the current study directly examined whether the size of

the self-association benefit in perceptual matching would be influenced by whether the competing stimuli were associated with 'favourable' or 'unfavourable' others.

In the present study participants associated three shape stimuli with the self, a stranger and either a politician for whom the participant holds a favourable opinion (Experiment 5) or a politician for whom the participant holds an unfavourable opinion (Experiment 6). A robust self-prioritization effect was found in both experiments; participants were significantly faster at matching shape stimuli to the self than to a politician. Also, participants were significantly faster to match stimuli to the politician than to the stranger. Importantly, cross-experiment comparisons revealed that holding a favourable or unfavourable opinion of the selected politician had little influence on the prioritization of the social stimuli for perceptual processing. There was no change to the magnitude of the self prioritization effect across experiments, as the speed and accuracy of self-match, politician-match and stranger-match responses did not significantly differ between experiments. Therefore, the data from this study provide no evidence to suggest that the social stimuli were prioritized for processing based on their 'likability' rather than their self-relevance.

Of note, analysis of non-match trials revealed that participants were faster to respond that the self-shape had been paired with a non-matching social label if the label referred to an unfavourable other (the disliked politician) than a neutral other (the stranger). No such label effect was found when the politician label referred to a favourable other. This finding that participants are fast to distinguish the self from a disliked other is consistent with the findings of Chapter 5. In Experiment 3 (self vs. hero vs. villain) participants were significantly faster to make a non-match response if the self-shape was paired with a villain label compared to when it was paired with a hero label. In Experiment 4 (self vs. favourable president vs. unfavourable president), participants were faster to perform a non-match response if the self-shape was paired with the label 'Trump' (their least-liked President) than 'Obama' (their most-liked president). I propose that this pattern of results reflects a form of self-serving bias. That is, a cognitive bias that serves to maintain or enhances self-esteem (Forsyth, 2008).

It has been well demonstrated that people have an inherent drive to minimize their shortcomings and exaggerate their virtues (Alicke & Sedikides, 2009). For example, people construe or remember events in a way that presents themselves in the most favourable light – but only to an extent that is credible to oneself and to others (Alicke & Govorun, 2005; Alicke & Sedikides, 2009; Sedikides & Gregg, 2008). Therefore, within the context of the present study, it is possible that the participants were driven to disassociate themselves from disliked others. This, in turn, resulted in a speeding of non-match responses when the self-shape was paired with a label representing an unfavourable other.

The finding that participants were significantly faster to respond to the politician-associated stimuli over the stranger-associated stimuli, irrespective of whether they were liked or disliked, might be explained by a number of factors. For one, asking participants to choose a politician at the beginning of the task meant that additional emphasis was placed on the importance of the politician associations for the purpose of the experiment. Therefore, demand characteristics (an experimental artefact caused by participants unconsciously changing their behaviour to fit their interpretation of the experiment's purpose; Orne, 1962) may have resulted in biased responding to the politician-associated stimuli. However, it should be noted that self-associated stimuli were still prioritized over politician-associated stimuli, despite there being no particular emphasis placed on the self-associations at the beginning of the experiment.

Wade and Vickery (2017) specifically examined the role of demand characteristics in perceptual matching tasks. In their experiment participants matched shapes to the labels 'self', 'stranger' and 'other'; and particular emphasis was placed on the stranger-associations by priming participants to think of 'stranger danger'. Their data revealed that self-associated stimuli were prioritized over stranger-associated stimuli, irrespective of the threat priming. However, in their study, participants were no faster to match shapes to the label 'Stranger' than they were to the neutral label 'Other'. Wade and Vickery therefore argued that the self-association advantage was resistant to demand characteristics.

The advantage of politician-association over stranger-association could also be accounted for in terms of depth of processing (Craik & Lockhart, 1972; Craik & Tulving, 1975; Eysenck & Eysenck, 1979. Asking participants to generate and select a mental representation of a politician based on evaluative judgements of their likability, requires additional elaborative processing that the stranger stimulus was not subject to. Of relevance, one of the key illustrations of the mnemonic benefit of elaborative processing is the 'self-reference effect' (see Chapter 1.2 for an overview of self-reference effects in memory). It is thought that the processing of stimuli related to the self will trigger widespread activation in that person's semantic network, creating stronger memory traces (Symons & Johnson, 1997). The self-association advantage in perceptual matching may reflect better memory of the association between a shape and the self. Likewise, the elaborative processing of the politician might itself lead to superior matching performance relative to the stranger.

Another factor to take in to account is that the politician (relative to stranger) has some degree of self-relevance because the politician was chosen by the participant. Therefore, the politician-associated stimuli may have been prioritized over the stranger-associated shape due to being encoded in reference to the self (i.e., the politician that *I* like or the politician *I* dislike). Previous perceptual matching studies have found that novel stimuli representing individuals associated with the self (i.e., the participant's mother or best friend) are prioritized for processing over stimuli associated with unfamiliar others (Sui et al., 2012). Therefore, evidence suggests that the prioritization of social stimuli in perceptual matching is driven by personal significance rather than likability.

In sum, the experiments presented in this chapter demonstrate that stimuli associated with the self are prioritized over stimuli associated with self-selected others. Of importance, the size of the self-association benefit remained stable irrespective of whether the other individuals were liked or disliked by the participant. Therefore, in conclusion, across four studies in total I found no evidence to suggest that the social stimuli were prioritized for processing based on the subjective 'likability' of the individuals.

7 TASK CO-REPRESENTATION

7.1 Abstract

Successful task sharing typically requires co-actors to simulate and predict each other's intentions and action plans. Research suggests that task co-representation occurs even in division-of-labour scenarios in which no interpersonal coordination is required. I investigated whether co-actors performing independent binary-choice responses to a mutually attended stimulus take into account each other's stimulus-response rules. In my novel paradigm, pairs of participants were required to simultaneously 'accept' or 'reject' a multi-feature stimulus based on whether it met their individually assigned target requirements. When participants were aware of their partner's task set they were faster to accept a stimulus if their co-actor's target colour was also present in the stimulus and faster to reject the stimulus if their co-actor's target colour was absent (Experiment 7). When participant was unaware of the co-actor's task set no such response modulation was found (Experiment 8). Also, no task-interference was found when the participant was aware of their co-actor's task set but believed the co-actor to be performing the task in another room (Experiment 9). This novel task provides the first direct evidence that when performing tasks in shared environments, people not only represent *that* their co-actor is responding but also *how* and *why* (i.e., full stimulus-response mappings). Furthermore, activation within the co-actor's task representation interferes with action-selection within one's own task.

7.2 Background

The ability to reason about the beliefs and intentions of others is a fundamental component of most social interaction. In particular, to successfully coordinate actions with another individual it is often beneficial to form shared representations of how, when, and by whom a particular action is to be carried out. For example, when two people pass one another on a narrow pathway and they wish to prevent a collision, it is mutually beneficial for them to a have shared understanding about who will make way for the other. In this scenario both individuals must adjust some aspect of their own action plan in response to the presumed intentions and action plans of the other.

Recent research has suggested that co-actors take into account each other's task components even in complete division-of-labor scenarios, in which no interpersonal coordination is required (e.g., Atmaca, Sebanz, & Knoblich, 2011; Böckler, Knoblich, & Sebanz, 2012; Eskenazi, Doerrfeld, Logan, Knoblich, & Sebanz, 2013; Sebanz, Knoblich, & Prinz, 2003). As such, it has been proposed that people quasi-automatically form cognitive representations of what nearby actors are doing, irrespective of whether it is beneficial to their own task performance.

Typically, joint-task interference has been investigated by taking a standard binary choice interference task and dividing it between two individuals, so that they each perform a complementary component of the task. The most widely used paradigm adapted in this manner is the Simon task (Simon & Small, 1969). In the original Simon task a stimulus is presented on the left or right half of a display, and participants are required to produce a spatially defined response to non-spatial features of the stimulus (e.g., press the left button if the stimulus is red, press the right button if the stimulus is green). In the Joint Simon task (Sebanz et al., 2003) this binary choice is divided so that one participant presses a button in response to one stimulus feature (e.g., red) and their partner presses a different button in response to the alternative feature (e.g., green). In essence, the co-actors are performing two complementary tasks, wherein on any given trial, one actor responds and the other does not (i.e. withholds a response).

When participants perform the standard Simon task alone, responses are faster when stimulus and response locations are compatible, despite stimulus location being

irrelevant to the task (The Simon Effect/ SE, Simon & Rudell, 1967). It is thought that stimuli automatically activate spatially corresponding responses, causing conflict when stimulus and response locations mismatch. This Simon Effect typically disappears when participants are required to respond to the presence of one specific feature – a 'Go/No-Go' task (Sebanz et al., 2003). In this solo Go/No-Go version of the Simon task only one stimulus feature and one response location are mapped, thus stimulus-response compatibility becomes irrelevant.

Importantly, when the binary-response Simon task was distributed between two participants, so that they each perform complementary Go/NoGo responses, the SE returned – a phenomenon known as the joint Simon effect (JSE, Sebanz et al., 2003; Sebanz, Knoblich, & Prinz, 2005). That is, even though participants were only responsible for pressing one response button (as with the solo Go/NoGo task), responses were quicker when stimulus and response locations were compatible. Sebanz and colleagues proposed that participants in the joint task condition had cognitively represented not only their own stimulus-response (S-R) rules but also the co-actor's S-R rules. The representation of both self- and other- generated responses, along with their relative locations, reintroduces the stimulus-response compatibility effects responsible for the SE. This phenomenon is known as task co-representation and is assumed to be an unavoidable consequence of the sociallytuned brain (Knoblich & Sebanz, 2006; Sebanz, Bekkering, & Knoblich, 2006; Sebanz & Knoblich, 2009).

However, subsequent research has cast serious doubt over task co-representation as an explanation for the JSE. Notably, a reliable JSE is observed even when the 'co-actor' does not actively participate in the task (Dolk et al., 2011), and when co-actors have no knowledge of each other's task rules (Dolk, Hommel, Prinz, & Liepelt, 2013). Furthermore, it has been demonstrated that a JSE can be induced when the co-actor is replaced with a moving non-biological object, such as a metronome or a Japanese waving cat (Dolk et al., 2013; Dolk, Hommel, Prinz, & Liepelt, 2014). As it is impossible to 'co' represent anything with an inanimate object, it is unlikely that a purely social mechanism is behind such effects.

7.2.1 Non-social accounts of joint-task interference

A number of alternative explanations for the JSE have been proposed. The common rationale behind these accounts is that aspects of the shared environment, including the co-actor and their actions, shape representations of the participant's own task (Prinz, 2015). For example, the spatial coding account (Dittrich, Dolk, Rothe-Wulf, Klauer, & Prinz, 2013; Dittrich, Rothe, & Klauer, 2012; Guagnano, Rusconi, & Umiltà, 2010) suggests that participants use the co-actor as a spatial frame of reference from which they can represent their own action (i.e. "If my co-actor is to the right of me I press the button on the left"). The relative spatial coding of responses reintroduces the stimulus-response compatibility effect. Therefore, adding a co-actor to the task environment is sufficient to elicit a JSE, without the need for task co-representation.

In support of this account, studies have found that the JSE is ameliorated when co-actors perform their tasks outside of each other's peri-personal space, as the spatial reference frame becomes less salient (Guagnano et al., 2010; Sellaro, Treccani, Rubichi, & Cubelli, 2013). However, this finding has not always replicated (cf. Welsh et al., 2013). The spatial coding account might also explain why non-biological or inactive 'co-actors' can create a JSE. It is likely that any attention-grabbing object or event can be used as a salient reference point on which a participant can frame their own actions (Dolk et al., 2013, 2014).

Another non-social explanation of the JSE is the referential coding account (Dolk et al., 2013). Grounded in both ideomotor theory (James, 1890) and Theory of Event Coding (TEC; Hommel, 2009; Hommel, Müsseler, Aschersleben, & Prinz, 2001) this account proposes that joint task interference is a product of how self- and other-generated actions are cognitively represented. According to TEC all perceived events, whether they are self-generated or other-generated, are represented by their sensory consequences (such as the sound of a button clicking, or the visual perception of a

button being depressed). This common coding system creates a discrimination problem when self- and other- generated actions are perceptually similar. In order to overcome this problem, emphasis must be placed on any features of the action that can be used to identify the actor. In the joint Simon task the most obvious difference between self- and other- generated actions is their relative spatial locations. Therefore, participants deliberately code their own actions as either 'to the right' or 'to the left', reintroducing the dimensional overlap between stimulus and response that is responsible for the Simon effect.

However, it is important to note that joint-task inference has been found outside of spatial compatibility tasks (e.g., Atmaca, Sebanz, Prinz, & Knoblich, 2008; Böckler et al., 2012; Eskenazi et al., 2013). For example, Atmaca et al. (2011) found distributing the Eriksen flanker task (Eriksen & Eriksen, 1974) across pairs of participants increased the size of the Go/No-Go flanker effect. In a standard (binary-choice) flanker task participants produce a spatial response to the identity of letter targets appearing at the center of a display (e.g., a right response is associated with the letter H, and a left response to the letter K). The target is surrounded by flankers that should be ignored. Flankers can signal a compatible response to the target (e.g., HHHHH), an incompatible response to the target (e.g., KKHKK), or act as neutral distractors (UUHUU). Despite flankers being nominally task-irrelevant, responses are found to be faster on compatible trials and slower on incompatible trials, relative to neutral trials.

Atmaca et al. (2011) investigated joint task interference by asking participants to carry out Go/NoGo versions of the Flanker task either in isolation or alongside a co-actor performing the complementary task component. In the joint Flanker task, pairs of participants were assigned one target each (e.g., the letter H or K) and each participant responded only to the presence of their own target (a joint Go/No-go task). Therefore, on Go trials a target might be surrounded by compatible flankers (HHHHH), neutral flankers (UUHUU) or incompatible flankers that are also potential targets for the co-actor (KKHKK). Atmaca et al. found that incompatible flankers caused greater interference in the joint Go/No-Go condition than in the solo Go/NoGo condition (the Joint Flanker Effect, JFE). The authors proposed that participants had associated

their co-actors target letters with the alternative response, increasing response conflict when it appeared as a flanker.

However, it should be noted that incompatible flankers still produced response interference under solo conditions, albeit to a lesser extent than in the joint task. This finding suggests that participants had represented an alternative stimulus-response rule even when performing the task in social isolation. Dolk et al. (2014) proposed that the presence of a co-actor does not induce the representation of an alternative rule, but instead draws attention to it. Furthermore, Dolk and colleagues demonstrated that, as with the JSE, a JFE can be provoked by the presence of a Japanese waving cat. Thus, the presence of the co-actor is not necessary for the representation of an alternative stimulus-response rule, but the presence of a non-biological moving object can strengthen interference caused by those alternative rules.

Alternatively, Wenke et al. (2011) suggested that agent identification (i.e. determining whose turn it is to make a response) may be the primary source of interference in joint Go/NoGo tasks. Instead of co-representing what the co-actor is supposed to be doing (i.e. S-R mappings), participants may simply be representing that another actor is responsible for half the responses, and when it is their turn to respond. Conflict relating to agent identification is especially applicable to tasks in which co-actors perform complementary responses, as stimulus-response compatibility is completely confounded with agent-response compatibility.

Wenke and colleagues provided evidence for this actor co-representation account of joint-task interference through a series of experiments using joint binary-choice tasks (Wenke et al., 2011). In Experiment 1 of their study co-actors took turns to make binary choice responses to the colour of a central target flanked by distractors. For example, one participant made a left key press if a central target was blue and a right key press if it was red, while the other participant made a left key press if a central target was yellow or a right key press if it was green. Responses were generally faster when the target was flanked by distractors of their own target colours (i.e., an own flanker advantage). However, when flankers were presented in their co-actor's target colours,

responses were not sensitive to whether flankers signalled a compatible or incompatible response to the target. This suggests that while participants had cognitively represented the stimulus conditions under which the co-actor should respond, they had not represented how they should respond (i.e., not full stimulus-response mappings).

7.3 Study Overview

As discussed above, there is substantial doubt as to whether joint-task interference is a result of automatic task co-representation, or whether non-social elements of the shared task environment are responsible. In the present study I present a novel paradigm that allows direct insight into what is being cognitively represented in joint tasks. In this new task, pairs of participants were asked to independently accept or reject the same multi-feature stimulus on the basis of whether it contained their individually assigned set of target features. Therefore, rather than taking turns to perform complementary Go/NoGo responses, each participant performed binary-choice responses on every trial.

Unlike the joint Go/NoGo tasks that have previously been used to investigate participants' awareness of when a co-actor should respond, this novel paradigm can be used to investigate awareness of how their co-actor should respond (i.e., whether they should accept or reject the stimulus) and why (i.e., which of their target features are present or absent). An added advantage is that there are no significant changes to the task environment across conditions: the number of actors, response alternatives, executed actions (and their relative locations) remained consistent within the experiment. As a result, the only two potential sources of interference are: 1) direct observation of the co-actor's actions, and 2) awareness of the co-actor's task rules. Across three experiments I investigate the relative contributions of these two forms of interference.

In Experiment 7 pairs of co-actors, seated side-by-side, were made aware of each other's task rules before performing independent binary-response tasks. Therefore, participants knew the identities of their co-actor's target features and could observe how their co-actor was responding. In Experiment 8 co-actors were able observe their partner's responses but were unaware of their partner's task set. The purpose of this experiment was to investigate how the observation of a co-actor's responses contributes to task interference. Finally, Experiment 9 investigated task interference when participants were aware of their partner's task set but performed their tasks in separate locations. The aim of this experiment was to investigate whether task interference occurs without visual or auditory feedback about the co-actor's responses.

7.4 The task co-representation paradigm

The aim of the study was to investigate whether having an awareness of a co-actor's binary-choice stimulus-response rules would influence action selection within one's own binary-choice task. To do so, I developed a novel paradigm in which pairs of participants searched the same multi-feature stimulus for their own set of target features. The stimulus consisted of a letter or the alphabet surrounded by two differently coloured squares (each presented in either red, blue, green, or yellow). Participants were required to 'accept' the stimulus (by pressing one response button) if the following conditions were met: the letter was a vowel AND either one of the two squares were presented in their assigned target colour. If either or both of these conditions were not met, then the participant must 'reject' the stimulus by pressing the other response button.

Importantly, co-actors were assigned *different* target colours. As the stimulus consisted of two squares, it was possible that both co-actor's target colours could be present within the same stimulus. Accordingly, the co-actor's tasks were not complementary: the presence/absence of one target colour could not be used to

determine the presence/absence of the other; and observing the action of the co-actor could not be used to determine one's own course of action. Any cross-task interference outside of that which can be accounted for by action observation would support the claim that people have an automatic tendency to cognitively represent the tasks of others, irrespective of whether it provides any benefit to the self (Sebanz et al., 2003). Therefore, rather than examining the effect of action compatibility between co-actor's (i.e., whether co-actors were required to respond in the same way to a specific stimulus) analysis focused on whether the presence of the co-actor's target colour within the stimulus influenced the selection of an accept or reject response.

In order to determine the correct response to a stimulus (i.e., either an accept or reject response), each co-actor must first code the individual stimulus features as 'targets' or 'non-targets' in accordance with their own task rules. Then, the following logical conjunction can be computed: IF the letter target AND the colour target are present within the stimulus, THEN perform an accept response, ELSE perform a reject response. However, if co-actors automatically form a shared representation of both tasks, then the colour coded as a 'target' in respect to one task would be coded as a 'non-target' in respect to the other. Therefore, selecting the correct course of action would require a participant to successfully identify and then prioritise the representation of their own task over the representation of their co-actor's task.

If the task co-representation account is correct in that individuals have an automatic tendency to represent the stimulus-response rules of nearby others, then it is possible that the presence of the co-actor's target colour within the stimulus will interfere with participant's performance. These interference effects could manifest as either:

1. General slowing: the co-actor's target feature acts as a salient task-irrelevant distractor, slowing responses across the board (irrespective of whether the participant is accepting or rejecting the stimulus).

2. Modulation: If co-actors co-represent full stimulus-response mappings then the co-actor's target colour would prime activation of an accept response within the shared representation. This task-irrelevant priming will promote the correct response when the participant is required to accept the stimulus (speeding responses), but promote the incorrect response when the participant is required to reject the stimulus (slowing responses).

7.5 Experiment 7: Same Room, Shared Task Instructions

In Experiment 7 pairs of co-actors sat side-by-side and independently responded to a mutually-attended stimulus. Each co-actor was given their own set of task instructions. Importantly, the co-actors were explicitly informed of each other's stimulus-response rules. Therefore, participants were aware that the co-actor's target colour differed from their own. The purpose of Experiment 7 was to determine whether the presence of the co-actor's target colour within the stimulus would influence the speed at which participants either accepted or rejected the same stimulus in accordance with their own task requirements.

7.5.1 Method

7.5.1.1 Participants

Thirty University of Warwick students (15 pairs, aged 18-44 years; M = 20 years, 8 male) participated in exchange for course credit or payment (£4). All participants reported that they had never met their task partner prior to the start of the experiment. All participants had normal colour vision and normal or corrected-to-normal visual

acuity. Participants gave informed consent prior to participation, in accordance with procedures approved by a local ethics committee.

7.5.1.2 Apparatus and stimuli

Displays were generated and responses recorded using custom programs running on a PC attached to a 19" LCD monitor running at a resolution of 1440 × 900 and 60Hz screen refresh rate. Stimuli were presented on a uniform grey background (RGB values = 128,128,128).

Each stimulus consisted of three overlapping components: a letter of the alphabet, and two square outlines (see Figure 22). The letter component was presented at the centre of the display, in black (RGB values = 0, 0, 0) upper case, Times New Roman typeface; subtending approximately 3° x 3°. On half of trials the letter was randomly selected with replacement from one of six vowels of the English/Latin alphabet (i.e., A, E, I, O, U). On the remaining half of the trials the letter was randomly selected with replacement from one of 20 consonants of the English/Latin alphabet (B, C, D, F, G, H, J, K, L, M, N, P, Q, R, S, T, V, W, X, Z). The letter Y was not included to avoid confusion, as it is considered a 'semivowel' within many languages (Maddieson & Emmorey, 1985). The letter was surrounded by two concentric square outlines of equal dimensions, one oriented at 0 degrees and the other at 45 degrees (Figure 22). Both squares subtended 8° in edge length. Each square could be presented in red (RGB values = 255, 0, 0), green (RGB values = 0, 255, 0), blue (RGB values = 0, 0, 255), or yellow (RGB values = 255, 255, 0); but there were never two squares of the same colour within a single stimulus.

Participants responded via USB hand-held PC-gamepads. Two buttons located equidistant from the centre of the gamepad (one on the top left and one on the top right) were used as response buttons. One button signified an 'Accept' response and the other a 'Reject' response, with the assignment counterbalanced across participants.

Labels of 'Accept' and 'Reject' were placed under the appropriate buttons to denote their assignment. Buttons were presses using their right or left index fingers.

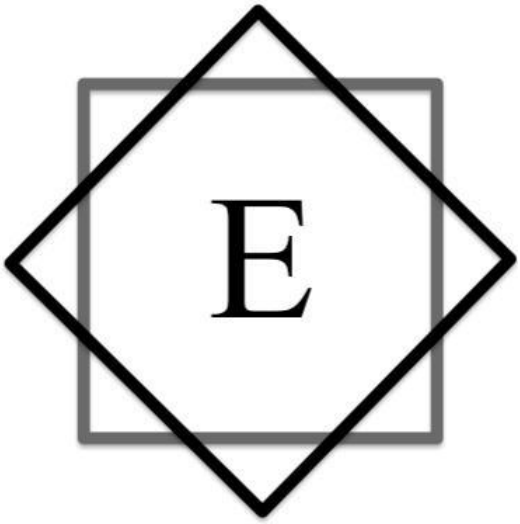

Figure 22. Example stimulus consisting of two differently coloured outline boxes and a central letter. The boxes could be red, green, blue or yellow, and the letter was either a vowel or a consonant. Targets were defined as containing a vowel and a specific colour in either box, e.g., a vowel and the presence of a red box.

7.5.1.3 Procedure

Two participants sat side-by-side in front of a shared computer monitor. The experimenter read aloud both participants' task instructions, so that each participant was aware of their own task set *and* the task set of their co-actor. For example, participants would hear the following instructions (emphasis in bold):

"*In each trial of the experiment you will see a letter surrounded by two overlapping squares. The squares will be coloured red, blue, green or yellow, each one a different colour. Participant 1 [on the left] press the 'accept' button on your controller if the letter is a vowel (A, E, I, O, U)* ***and*** *there*

is a blue square present. If the letter is not a vowel and/or there is no blue square present, you must press the 'reject' button on your controller. Participant 2 [on the right] you must press the 'Accept' response if the letter is a vowel ***and*** *there is a red square present. If the letter is not a vowel and/or there is no red square, press the 'Reject' response on your controller."*

Each trial began with the presentation of a central fixation cross for 1000ms. This was followed by the presentation of a single stimulus (a letter surrounded by two squares), which remained visible until both participants had responded or 2000ms had elapsed. Each participant made a response by pressing one of two buttons on their individual gamepads. The assignment of each button to indicate either an 'Accept' response or a 'Reject' response was counterbalanced across participants. Participants pressed their 'Accept' response button if the stimulus met *both* of their target criteria (vowel + target colour), or their 'Reject' button if the stimulus failed to meet either or both target criteria. Once both participants had made a response the trial was immediately ended. However, on practice trials only, response feedback of 'correct response' or 'incorrect response' was presented on the display at the end of each trial for 4000ms. For the participant seated to the left of centre, feedback was presented on the left half of the display. For the participant seated to the right of centre, feedback was presented on the right half of the display.

Participants first completed a practice block of 10 random trials. The full experiment consisted of 10 blocks of 64 trials. Within each block there were eight repetitions of eight trial types, defined by which of the actor's target features were present (colour + vowel, colour only, vowel only, none) and which of the co-actor's target features were present (colour + vowel, colour only, vowel only, none). Table 8 depicts each of the 8 possible trial types. Trial order was randomised within each block.

Table 8. Trial types defined by which of the actor's and co-actor's target features were present in the stimulus, and the required response from each participant. A tick indicates the presence of a feature or action, whereas a cross indicates the absence of a feature or action.

Trial Type	Letter is a vowel?	Actor's Colour present?	Co-actor's colour present?	Actor to ACCEPT stimulus	Co-actor to ACCEPT stimulus	Actor to REJECT stimulus	Co-actor to REJECT stimulus
1	✓	✓	✓	✓	✓	✗	✗
2	✓	✓	✗	✓	✗	✗	✓
3	✓	✗	✓	✗	✓	✓	✗
4	✓	✗	✗	✗	✗	✓	✓
5	✗	✓	✓	✗	✗	✓	✓
6	✗	✓	✗	✗	✗	✓	✓
7	✗	✗	✓	✗	✗	✓	✓
8	✗	✗	✗	✗	✗	✓	✓

7.5.2 Results

Outlier RTs (<200ms or >2000ms) were discarded (<1% of trials). If participants performed at or below 50% accuracy in any condition their data were discarded (one participant). To examine the specific effect that the presence of the co-actor's target colour had on participant's responses, trials were first grouped into four conditions depending on which of the participant's target features were present in the stimulus ('colour and vowel', 'colour only', 'vowel only', 'none'). For each of the four conditions, responses were compared when the co-actor's target colour was present and when the co-actor's colour was absent. Table 9 presents the error rates (proportion correct) for each condition. Overall, accuracy was very high (mean proportion correct =.93) and there was no evidence of a speed-accuracy trade-off. Figure 23 displays the

mean RTs for each condition, as well as mean difference in RTs between trials when the co-actor's target colour was present and when it was absent.

Table 9. Accuracy as a Function of which of the Participant's Target Features are Present in the Stimulus, and the Presence or Absence of the Co-actor's Target Colour within the Same Stimulus, for Experiment 7

Own Target features present	Correct response type	Accuracy	
		Co-actor's colour present	Co-actor's colour absent
Colour and Vowel	Accept	.88 (.09)	.86 (.10)
Colour only	Reject	.90 (.06)	.92 (.05)
Vowel only	Reject	.95 (.07)	.98 (.03)
None	Reject	.98 (.04)	.98 (.03)

Note. Accuracy = proportion correct. Standard deviations appear within parentheses.

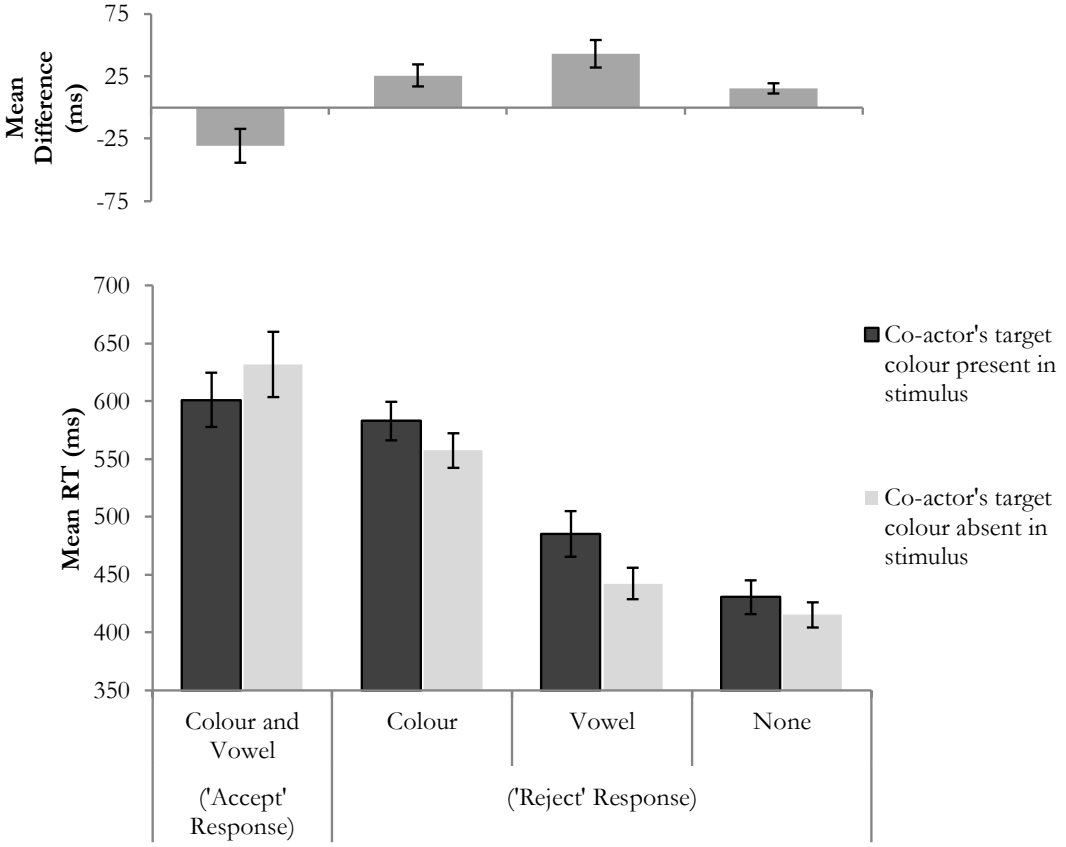

Figure 23. Mean RTs (bottom) and RT differences between conditions (top) for Experiment 7. The x axis denotes which of the participant's own target features were present within the stimulus, with the correct response in brackets. Error bars denote ± 1 standard error of the mean.

7.5.2.1 RTs

Due to the non-orthogonal design of the experiment, analysis consisted of a number of specific within-subjects pair-wise comparisons. For each of the four trial types ('colour and vowel', 'colour only', 'vowel only, 'none') RTs were compared when the co-actor's target colour was present within the stimulus to when the co-actor's target

colour was absent. All reported p values were adjusted post hoc using the holm-bonferroni method (Holm, 1979).

When the stimulus met both target requirements ('colour and vowel' condition), requiring the participant to make an accept response, RTs were significantly faster when the co-actor's target colour was also present within the stimulus, $t(28) = -2.26$, $p = .03$, $d = -.42$. When participants were required to make a reject response because one or both or their target requirements had not been met by the stimulus, RTs were significantly slowed by the presence of the co-actor's colour within the stimulus. This was true for the 'none' condition, the 'colour only' condition, and the 'vowel only' condition; $t(28) = 3.67$, $p = .004$, $d = .68$, $t(28) = 2.92$, $p = .01$, $d = .54$, $t(28) = 3.84$, $p = .004$, $d = .71$, respectively.

7.5.2.2 Accuracy

For each condition ('colour and vowel', 'colour only', 'vowel only, 'none') response accuracy (proportion of correct responses) was compared when the co-actor's target colour was present in the stimulus to when it was absent. When the stimulus contained the participant's own conjunction of target features ('colour and vowel' condition), requiring the participant to accept the stimulus, the presence of the co-actor's target colour within the stimulus was not found to influence response accuracy, $t(28) = 1.01$, $p = .64$, $d = .19$. When participants were required to reject the stimulus because the letter was a consonant ('colour only' condition), the target colour was absent (the 'vowel only condition'), or the letter was a consonant *and* the target colour was absent ('none' condition), the presence of the co-actor's target colour within the stimulus had no significant influence on the accuracy of responses; $t(28) = -1.53$, $p = .42$, $d = -.28$, $t(28) = -2.48$, $p = .08$, $d = .46$ and $t(28) = -.99$, $p = .64$, $d = -.18$, respectively.

7.5.3 Discussion

The results of this experiment provide clear evidence of cross-task interference between co-actors, despite the independence of their tasks. In particular, participants were *faster* to 'accept' a multi-feature stimulus for meeting their own target requirements if their co-actor's target colour was present in the stimulus, but *slower* to 'reject' a stimulus for not meeting their target requirements if the co-actor's target colour was present in the stimulus. Of particular note, interference could not be fully accounted for by direct action observation (i.e., observing whether the co-actor was performing a compatible or incompatible response). When the letter component of the stimulus was not a vowel (the 'colour only' and 'none' conditions), requiring both co-actors to perform a reject response, RTs were still slower when the co-actor's target colour was present within the stimulus.

This pattern of cross-task interference is consistent with participants holding two simultaneous representations of the task: one of their own stimulus-response mappings, and one of the stimulus-response mappings of their co-actor. An inability to fully supress activation within the representation of the co-actor's task resulted in the priming of an accept response whenever the participant perceived the co-actor's target colour. This extraneous priming facilitated the execution of the correct action when participants were required by their own stimulus-response rules to accept the stimulus, speeding responses. However, when the participant was required to reject the stimulus, this task-irrelevant priming required top-down suppression in order to reduce action-selection conflict, slowing responses.

The novel finding from Experiment 7 was that having an explicit awareness of a co-actor's task set *modulates* own performance on a binary-choice task. This is direct evidence of task co-representation outside of complementary turn-taking scenarios. Furthermore, the results suggest that participants had mapped the co-actor's target features to a *specific* response. The results of this study show that co-representation is not

limited to direct stimulus-response mappings, but also more complex binary-choice tasks.

7.6 Experiment 8: Same Room, Private Task Instructions

Having an explicit awareness of a co-actor's binary-choice stimulus-response rules has been found to interfere with performance on one's own binary-choice task (Experiment 7). In Experiment 8, co-actors performed their independent tasks side-by-side, but participants were not made aware of their co-actor's task rules. The purpose of this experiment was to determine whether cross-task interference occurs even in the absence of awareness of the co-actor's task rules.

As participants held no prior knowledge of the co-actor's target colour, it was expected that the presence or absence of the co-actor's target colour within the stimulus would have no influence on own-task performance. The only possible source of cross-task interference is the ability to witness how the co-actor is responding. According to an extended form of the ideomotor principle (Greenwald, 1970; Jeannerod, 1999; Prinz, 1997), observing an action performed by another individual activates the same representational structures as those involved in the planning and execution of one's own action. Behavioural evidence in favour of automatic action co-representation from studies showing that the performance of an action is facilitated when concurrently observing another person performing the same action, but impaired when observing another person performing a different action (Brass et al., 2009; Spengler, von Cramon, et al., 2010). In line with this action co-representation account, it is possible that participants would be faster to perform a response when the co-actor was performing the same type of response (i.e., both accepting or both rejecting the stimulus). However, this would only occur if participants were actively observing their partner's actions– something that is not required by the task, nor beneficial to the participant's own performance.

Although participants were not made explicitly aware of their co-actor's task rules at the beginning of the task, there exists the possibility that participants infer their co-actor's task rules from observing their actions. That is, when a specific observed action occurs under consistent conditions, such as there always being a blue square present when the co-actor performs an accept response, it is possible that the participant deduces that the co-actor's target colour is blue. Then, once the co-actor's target colour becomes co-represented, a similar pattern of interference would be expected as that of Experiment 7 (i.e., that the co-actor's target colour speeds accept response but slows reject responses). However, participants each responded on handheld response panels and it would be difficult to observe both the monitor and the co-actor's hands. If participants exerted the cognitive effort to work out and represent the co-actor's task rules despite it being unnecessary for their own task, this would be testament to the drive to understand and predict the behaviours of others.

7.6.1 Method

7.6.1.1 Participants

Thirty participants (15 pairs, aged 18-27 years; M = 22 years 10 male) took part participated in exchange for course credit or payment (£4). Co-actors reported that they had not met each other prior to the experiment. All participants had normal colour vision and normal or corrected-to-normal visual acuity. Participants gave informed consent prior to participation, in accordance with procedures approved by a local ethics committee.

7.6.1.2 Apparatus and Stimuli

The apparatus, stimuli, and task were the same as that of Experiment 7. The key deviation was that each participant was presented with written task instructions on paper prior to the start of the task.

7.6.1.3 Procedure

The general procedure was the same as that of Experiment 7, except that each participant was given the following written task instructions privately and were told not to share their task instructions with their co-actor:

In each trial of the experiment you will see a letter surrounded by two overlapping squares. The squares will be coloured red, blue, green or yellow; each one a different colour. Press the 'accept' button on your controller only if both of the following are true: the letter at the centre of the stimulus is a vowel (A, E, I, O, U) ***and*** *either of the two squares if [insert target colour]. If the letter is not a vowel* ***and/or*** *there is no [target colour] square present, you must press the 'reject' button on your controller*

As participants sat side-by-side, in front of a mutually attended computer display, and with no physical divide between them, it was possible for co-actors to witness each other's responses (i.e., whether they were pressing the accept or reject button). However, as each participant responded on a handheld control panel it was difficult (but not impossible) to determine whether the co-actor was performing an accept or reject response.

7.6.2 Results

The methods of data analysis and initial exclusion criteria were the same as those described in Experiment 7. RTs below 200ms or above 2000ms were discarded (<1% of trials). Data from one participant were discarded (<50% correct responses in one or more condition). Accuracy data are presented in Table 10. Figure 24 displays the error-corrected mean RTs for each condition, as well as mean difference in RTs between trials when the co-actor's target colour was present and when it was absent.

Table 10. Accuracy as a Function of which of the Participant's Target Features are Present in the Stimulus, and the Presence or Absence of the Co-actor's Target Colour within the Same Stimulus, For Experiment 8.

Own Target features present	Correct response type	Accuracy	
		Co-actor's colour present	Co-actor's colour absent
Colour and Vowel	Accept	.88 (.08)	.87 (.09)
Colour only	Reject	.94 (.04)	.94 (.04)
Vowel only	Reject	.96 (.05)	.97 (.03)
None	Reject	.99 (.02)	.98 (.02)

Note. Accuracy = proportion correct. Standard deviations appear within parentheses.

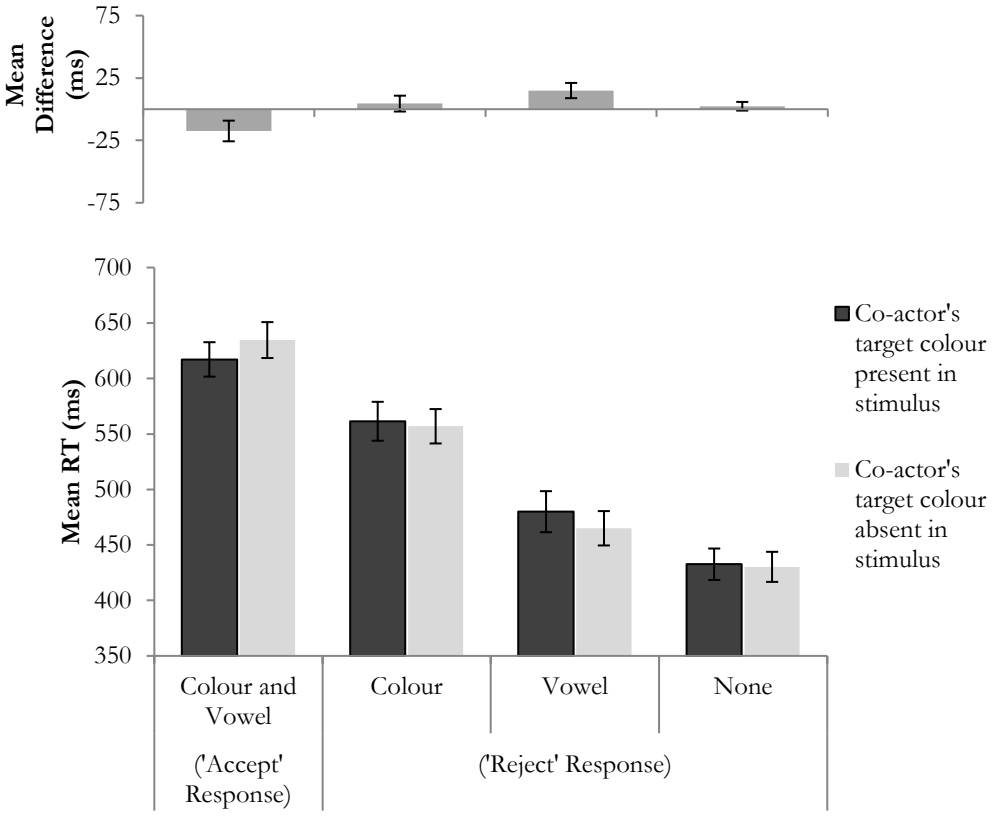

Figure 24. Mean RTs (bottom) and RT differences between conditions (top) for Experiment 8. The x axis denotes which of the participant's own target features were present within the stimulus, with the correct response in brackets. Error bars denote ± 1 standard error of the mean.

7.6.2.1 RTs

All pairwise comparisons were subject to Holm-Bonferroni corrections for multiplicity (Holm, 1979). When the participant was required to accept the stimulus for containing both their target features ('colour and vowel' condition), the presence of the co-actor's target feature had no significant influence on RTs, $t(28) = -2.32$, $p = .03$, $d = -.43$. When the participant was required to reject the stimulus because the central letter was not a vowel but their target colour was present ('colour only' condition), the target colour had no significant influence on RTs, $t(28) = 1.48$, $p = .15$, $d = .28$. Also, the presence of the co-actor's colour had no significant influence on reject responses when

both the letter and vowel conditions were not met ('none' condition); $t(28) = 1.36$, $p = .19$, $d = .25$. However, when the stimulus contained a vowel but not the participant's target colour ('vowel only' condition) participants were significantly slower to reject the stimulus if it contained the co-actor's target colour, $t(28) = 3.18$, $p = .03$, $d = .58$.

7.6.2.2 Accuracy

For all conditions ('colour and vowel', 'colour only', 'vowel only' and 'none' conditions), the presence of the co-actor's target colour within the stimulus had no influence on the accuracy of responses (all ts < 1.5, all ps > .14)

7.6.3 Discussion

In Experiment 8 pairs of participants performed independent binary-choice tasks in response to the same multi-feature stimulus. Unlike Experiment 7, participants were not explicitly informed of their co-actor's task instructions. Therefore, it was not expected that the presence of the co-actor's colour within the stimulus would have any effect on participants' performance.

In three conditions ('colour and vowel', 'colour only' and 'none' conditions) the presence or absence of the co-actor's target colour had no influence of performance. However, in one of the four conditions (the 'vowel only' condition) the presence of the co-actor's target colour significantly slowed reject responses. This finding is unexpected, given that participants had no awareness of which of four possible colours their co-actor's target colour was.

Why participants were slower to reject a stimulus in the 'vowel only' condition when the co-actor's target colour was present, compared to when the co-actor's target colour was absent, is not apparent. One possible explanation is that observing a co-actor perform a different response to the self may have slowed response times. In the 'vowel only' condition the participant was required to make a *different* responses from the co-actor when the co-actor's target colour is present (i.e., the co-actor accepts the stimulus

and the participant rejects the stimulus) but the *same* response when the co-actor's target colour is absent (i.e., both the co-actor and the participant must reject the stimulus).

Studies investigating the link between action observation and action observation have typically employed "automatic imitation" or "visuomotor priming" paradigms (e.g., Brass et al., 2009; Genschow et al., 2017; Heyes, 2011; Hofree, Urgen, Winkielman, & Saygin, 2015). For example, when participants are asked to perform an action that is either compatible with an observed movement there is evidence of a facilitation effect. However, when asked to perform an action that is incompatible with the observed action, there is evidence of an interference effect (i.e., a slowing or disruption of performance). A similar response compatibility effect could be occurring in the present task.

However, there are a number of reasons why action imitation is an unlikely explanation for this finding. First, each participant made their responses on individual handheld controllers, making it difficult but not impossible for them to witness their co-actor's responses. Furthermore, participants were performing their own relatively demanding task, which required them to attend to a computer monitor and to their own response buttons. Therefore, actively attending to the co-actor's actions would be highly detrimental to own-task performance. Finally, and most importantly, response compatibility had no significant effect on RTs in the 'colour and vowel' condition. When participants were required to accept the stimulus for meeting their target requirements, the co-actor was also required to perform a compatible accept response if their own target colour was present in the stimulus, else they must make an incompatible reject response . In the 'colour and vowel' condition, however, the presence or absence of the co-actor's target colour had no influence on the speed at which the participant accepted the stimulus. Thus, it is unlikely that response compatibility between co-actor influenced responding.

It is difficult to account for why the presence of the co-actor's target colour would only influence performance in the 'vowel only condition' when participants had no awareness of what the co-actor's target colour was. It is important to note that the

presence and absence of the co-actor's target colour had no influence on performance in any of the other three conditions. This is in contrast with Experiment 7, in which the presence/absence of the co-actor's target colour influenced performance across all four conditions. In fact, having an awareness of the co-actor's task set was found to modulate own-task performance in a manner indicative of task co-representation. That is, participants were faster to accept a stimulus for meeting their own target requirements if the co-actor's target colour was also present, yet slower to reject a stimulus for not meeting target requirements if the co-actor's target was present (i.e., a clear modulation effect). No such modulation effect was found in the present experiment were participants had no awareness of their co-actor's task set.

7.7 Experiment 9: Different Rooms, Shared Task Instructions

In Experiment 9, pairs of participants were jointly informed of their own and each other's task rules but were then separated in to different cubicles to perform their independent tasks. The aim of this experiment was to determine whether participants take in to account the task rules of an absent co-actor when performing their own binary-choice response task. If participants encode an offline representation of the task that includes both their own task rules and their 'co-actor's' task rules, it is possible that cross-task interference will occur despite the absence of the co-actor within the task environment. Alternatively, task co-representation may only be sustained within a social environment.

Previous studies have indicated that co-representation occurs even in the absence of on-line auditory or visual feedback about the co-actor's actions (Milward et al., 2014; van der Wel, 2015; Vlainic et al., 2010). The findings of these studies suggest that, as a result of task co-representation, observing a stimulus is sufficient to trigger a representation of the co-actor's response. However, in these studies, co-actors performed their tasks within a shared environment, only their actions were shielded

from each other. The purpose of Experiment 9 was to determine whether task co-representation is a truly social phenomenon, inspired by the presence of the other person. To do so, in Experiment 7 participants were made aware of both their own and a co-actor's task set, but the co-actor was said to be performing their task in a separate room. If interference caused by task co-representation is triggered by the social environment, then the modulation of responses in relation to an alternative task set should not be found when the task is performed in isolation.

There is some evidence to suggest that task co-representation is a social effect entirely dependent on the presence of a co-acting person. For example, Böckler et al. (2012) found that when pairs of participants performed independent Navon tasks side-by-side, in which co-actors were required to either have a shared focus of attention or different foci of attention, performance was significantly inferior when the co-actor's foci of attention differed. This interference effect suggests that the co-actors took in to account the co-actor's expected focus of attention, despite being irrelevant to their ow task. However, the cross-task interference effect disappeared when the co-actors performed their individual tasks components in social isolation, despite having read the instructions specifying the co-actor's task.

If task co-representation is an automatic consequence of the social environment and must be triggered by the presence of a nearby actor, then it would be expected that the presence of the co-actor's target colour would have no influence on responses when the co-actor is absent. Alternatively, if co-actors formulate an offline co-representation of the shared task that is then consulted irrespective of whether the co-actor is present within the task environment, it would be expected that the presence of the co-actor's target colour would modulate performance in a similar manner to that of Experiment 7.

7.7.1 Method

7.7.1.1 Participants

Thirty participants (15 pairs, aged 19-29 years, *M* = 20; 10 male) took part in exchange for course credit or payment (£4). All participants reported that they had never met their task partner prior to the start of the experiment. All participants had normal colour vision and normal or corrected-to-normal visual acuity. Participants gave informed consent prior to participation, in accordance with procedures approved by a local ethics committee.

7.7.1.2 Apparatus and Stimuli

The apparatus, stimuli and task structure were the same as Experiments 7 and 8. However, rather than two participants responding to the same stimuli on a shared computer monitor, individual participants performed their tasks on separate computers in different rooms.

7.7.1.3 Procedure

The general procedure of Experiment 9 was the same as Experiment 7. However, in this version of the task the co-actors were introduced to each other and then separated in to different cubicles to perform their own tasks. However, before being separated the experimenter read aloud both co-actor's task instructions (in the same manner as reported in Experiment 7), so that participants were made explicitly aware of their own stimulus-response rules and the stimulus-response rules of their 'co-actor'.

7.7.2 Results

The methods of data analysis and initial exclusion criteria were the same as those described in Experiment 7 and Experiment 8. RTs below 200ms or above 2000ms were

discarded (<1% of trials). Data from one pair of co-actors (two participants) were discarded due to a recording error. Mean RTs and accuracy data are presented in Table 11. Figure 25 displays the error-corrected mean RTs for each condition, as well as mean difference in RTs between trials when the co-actor's target colour was present and when it was absent. All paired comparisons were subject to Holm-Bonferroni corrections for multiplicity (Holm, 1979).

Table 11. Accuracy as a Function of which of the Participant's Target Features are Present in the Stimulus, and the Presence or Absence of the Co-actor's Target Colour within the Same Stimulus, for Experiment 9.

Own Target features present	Correct response type	Accuracy	
		Co-actor's colour present	Co-actor's colour absent
Colour and Vowel	Accept	.93 (.08)	.92 (.06)
Colour only	Reject	.93 (.08)	.92 (.06)
Vowel only	Reject	.96 (.04)	.96 (.05)
None	Reject	.98 (.03)	.98 (.02)

Note. Accuracy = proportion correct. Standard deviations appear within parentheses.

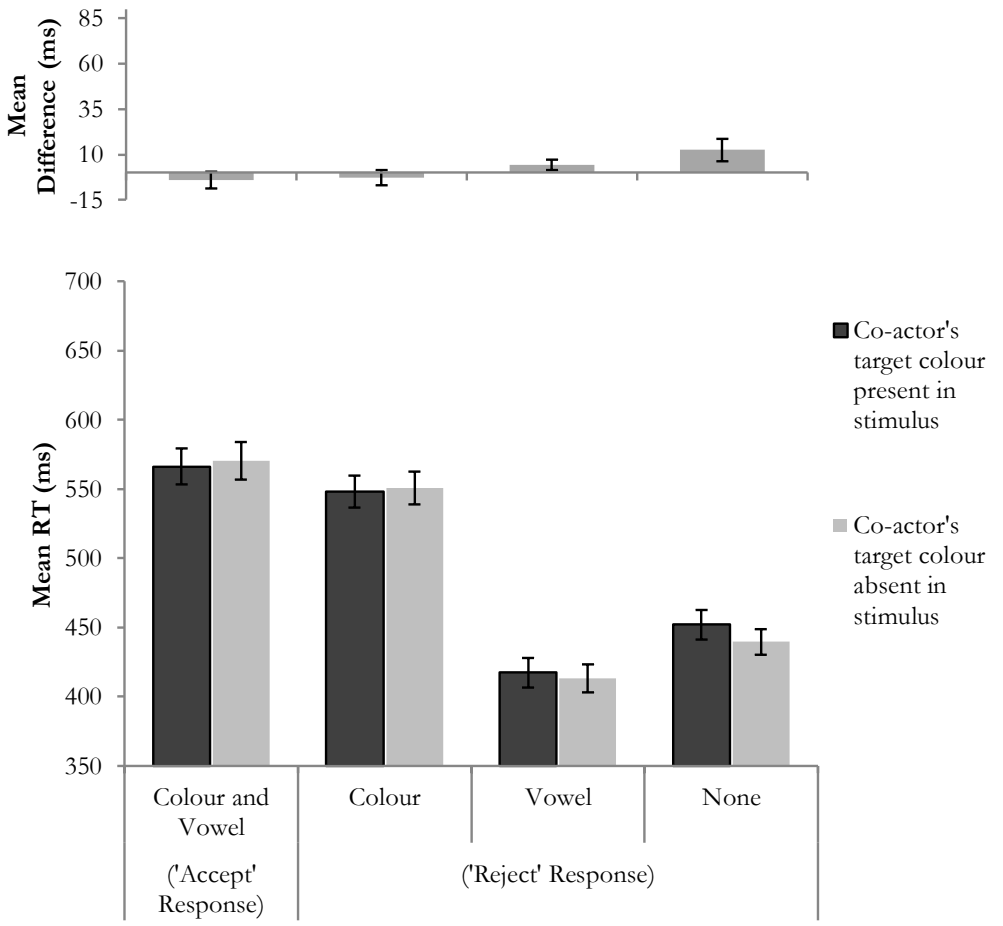

Figure 25. Mean RTs (bottom) and RT differences between conditions (top) for Experiment 9. The x axis denotes which of the participant's own target features were present within the stimulus, with the correct response in brackets. Error bars denote ± 1 standard error of the mean.

7.7.2.1 RTs

The presence of the co-actor's target colour in the stimulus had no influence on RTs in the 'colour and vowel', 'colour only', 'vowel only', or 'none' conditions; $t(27) = -1.15$, $p = .52$, $d = -.22$, $t(27) = -1.04$, $p = .52$, $d = -.19$ and $t(27) = 1.58$, $p = .39$, $d = .29$, , $t(27) = 2.46$, $p = .08$, $d = -.46$, respectively.

7.7.2.2 Accuracy

The presence of the co-actor's target feature within the stimulus had no significant influence on the accuracy of responding across any of the four conditions ('colour and vowel', 'colour only', 'vowel only', 'none'), all ts < 1.33, ps > .2.

7.7.3 Discussion

The aim of Experiment 9 was to determine whether participants take in to account the task rules of an absent co-actor when performing their own binary-choice response task in isolation. Overall, there was no evidence of cross-task interference. Participants' responses to the presence or absence of their own target features were unaffected by the presence or absence of the co-actor's target colour. Therefore, the co-actor's target colour was found to have no more influence on responding than other non-target colours within the stimulus. This pattern of results indicates that participants had not stored an offline representation of the co-actor's task set, or were able to successfully prioritise their own task set over their co-actor's task set when performing the task in isolation. In contrast, when co-actors were made aware of each other's task instructions AND performed their tasks side-by-side (Experiment 7) there was clear evidence of cross-task interference.

The results of Experiment 9 suggest that task co-representation may be a socially triggered phenomenon. That is, the automatic tendency to form and maintain a cognitive representation of another individual's task set is only triggered when performing a task alongside another physically present individual. However, in contrast to the present study, previous studies have found that the mere belief of acting together with another individual (real or virtual) is sufficient to trigger co-representation (Atmaca et al., 2011; e.g., Ruys & Aarts, 2010; Sebanz et al., 2005; Vlainic et al., 2010).

Alternatively, it may be that task co-representation can occur in the absence of the co-actor, but that the representation of the co-actor's stimulus-response mappings

become attenuated. Thus, for relatively simple tasks with direct stimulus-response mappings, observing the co-actor's target feature may be sufficient to activate a representation of the co-actor's response alternative when the co-actor is absent. However, in relatively complex tasks, in which stimulus features do not map directly to a single response (such as the present task), the co-actor's target colour was not enough to activate a co-actor's task representation.

It is also possible that participants did not believe that a co-actor was performing a task in a different room. No check was performed to determine the participant's beliefs about the co-actor. In a similar study conducted by Atmaca, Sebanz and Knoblich (2011), in which participants told that a co-actor was performing a task in a different room, just under 30% of participants reported afterwards that they doubted if the co-actor existed. Therefore, it is possible that an alternative task set will only be represented if the participant believes a co-actor to be performing the task.

What is important to note is that simply being aware of an alternative task set did not result in the pattern of cross-task interference found when two co-actors perform their tasks in a shared environment (Experiment 7). This supports the idea that task co-representation is a socially driven phenomenon that requires (1) an awareness of the co-actor's task set, and (2) a shared task environment.

7.8 Discussion

Previous research has indicated that co-acting individuals cognitively represent each other's tasks, even when interpersonal coordination is not required (e.g., Sebanz et al., 2003; 2005). While it has been demonstrated that participants are sensitive to the stimulus conditions under which a co-actor should perform an action (Atmaca et al., 2011, 2008; Sebanz et al., 2003), the current study is the first to provide direct evidence that co-actors are also sensitive to *how* their co-actor should react and *why*. In this study I asked pairs of participants to perform concurrent but independent binary-choice

responses to the same stimulus, in order to determine whether responses to the presence or absence of one's own target features were influenced by the presence or absence of the co-actor's target features.

In Experiment 7 pairs of participants were informed of both their own and their co-actor's task rules, before performing their respective tasks side-by-side. The data revealed that participants were faster to accept, but slower to reject, a stimulus when their co-actor's target colour was present. This pattern of results is evidence that participants were not only aware of the identity of their co-actor's target colour but had mapped it to a particular response (i.e., an accept response). Of note, the presence of the co-actor's target colour slowed reject responses irrespective of whether the co-actor was making an accept or reject response. This indicates that interference was caused by the presence of the co-actor's target colour but not the response demanded by their full task set. I propose that while participants had co-represented individual feature-response mappings, they had not combined these mappings to the extent that they could simulate the co-actor's correct response to a stimulus.

In Experiment 8 co-actors performed their tasks together but were unaware of each other's task rules. Overall, the presence or absence of the co-actor's target colour within the stimulus had no effect on performance. This is unsurprising, given that participants were unaware of which target colour had been assigned to their co-actor. There was the unexpected finding that the presence of the co-actor's target colour slowed reject responses in one condition (the 'vowel only' condition), which is difficult to explain. One possibility is that participants were slower to respond when they witnessed their co-actor perform a different response to their own (Wel & Fu, 2015). However, as the task was relatively demanding, and responses were made on hand-held controllers, it seems unlikely that the participants were observing each other's actions.

In the final experiment of this chapter (Experiment 9) pairs of participants were informed of their own and each other's task sets, before being separated in to different rooms to perform their tasks in isolation. Participants' responses to the presence or absence of their own target features was unaffected by the presence of the co-actor's

target colour within the stimulus. Overall, the findings of this experiment suggest that task co-representation is triggered and reinforced by the presence of the co-actor in the task environment. Therefore, when the co-actor was absent from the task environment cross-task interference no longer occurred, despite having an initial awareness of the co-actor's task set.

Taken together, findings from the three experiments identify two sources of joint-task interference: 1) the offline representation of a co-actor's task, and 2) the online monitoring of a co-actor's actions. It is important to consider why individuals remain tuned-in to others when doing so is detrimental to their own task performance. Sebanz et al. (2005) proposed that the cognitive requirements of joint action have shaped the cognitive system. In order for people to work together to achieve common goals, they must be able to share and represent each other's intentions and action plans effectively. Although the current task did not require interpersonal coordination, it may be that the cognitive system is set up to be prepared for joint action whenever a 'co-actor' is perceived.

In a more general sense, there is a significant adaptive benefit to being able to quickly simulate and predict the intentions of others; both in terms of cooperation and threat avoidance. As such, the benefits of automatic task co-representation are likely to outweigh the costs over time. In the real world, the social environment is usually less regimented; tasks get disrupted, rules change, and people must be able to modify their action plans. It is therefore valuable to not only be aware of what is happening in real time, but to theorize about what people might do in the near future.

The mechanisms behind task co-representation have been subject to debate over the last decade. One proposed theory is that the observation of a co-actor's actions activates corresponding perceptual and motor programs in the observer (Blakemore & Decety, 2001; Prinz, 1997; Vesper et al., 2017; Wilson & Knoblich, 2005). However, my results show that co-representation is not limited to direct perception-action links. It has also been proposed that difficulties in self-other distinction may play a significant role in unintentional co-representation (Milward & Sebanz, 2016). A study looking at the

relationship between task co-representation and cognitive executive functions in children has found a positive relationship between performance on a Theory of Mind task and protection against interference from a partner's task set (Milward, Kita, & Apperly, 2014). Alternatively, it may be that unintentional task co-representation and task monitoring are an overflow of the socially tuned mind (Sebanz et al., 2005). Given the overall benefits of knowing what nearby actors are doing, it is plausible that we have resources specifically reserved for the processing of other people's actions and intentions.

In sum, this study provides direct evidence that people in shared environments cognitively represent and continuously monitor the intentions and actions of a present co-actor, even when their task does not require interpersonal coordination. Most importantly, this study demonstrates for the first time that participants not only represent that a co-actor is responding, but how they should respond in respect to different stimulus conditions.

8 Discussion

The aim of this final chapter is to provide an integrated discussion of the findings presented in Chapters 4 to 7. Three main themes based around the topic of self-prioritisation were addressed throughout these experimental chapters:

1. The prioritisation of self-associated information over information associated with other individuals present or absent from the task environment.
2. The prioritisation of self-associated information over information associated with liked and disliked others.
3. The prioritisation of one's own task over the task of a present or an absent co-actor.

First, this chapter will provide an overview of these three lines of investigation. Second, I will discuss the potential applications and future directions for this field of research.

8.1 The prioritization of self-associated information over information associated with other present or absent individuals

Previous studies of self-biases in perceptual processing have involved single participants responding to self- and other- associated stimuli in social isolation (Humphreys & Sui, 2015a; Sui et al., 2012; Sui & Humphreys, 2013b). It was therefore unknown to what extent self-bias is modulated by the shared social environment, or whether the ability to 'tag' information to a physical body enhances social salience.

In two perceptual matching experiments, participants learned associations between neutral geometric shapes and three social labels (You, Stranger, or Partner), and following this, judged whether a series of presented shape-label pairings matched or mismatched. In Experiment 1 participants performed the task in isolation and were told that the label 'Partner' referred to a co-actor performing the same task in another room. In Experiment 2 the label 'Partner' referred specifically to a co-actor performing a perceptual matching task alongside the participant.

Overall, the experiments in Chapter 4 revealed five key findings:

1. Self-associated stimuli were prioritised for processing over stimuli associated with all other individuals (present or absent).
2. The size of the self-association benefit was impervious to whether the task was carried out in isolation or within a shared environment.
3. The size of the self-association benefit was impervious to whether competing stimuli were associated with a physically present co-actor or an imagined individual.
4. Visual stimuli associated with a co-actor were not prioritised for processing over stimuli associated with an imagined stranger, irrespective of whether the co-actor was present or imagined.
5. Participants were slower to indicate that a self-associated stimulus had been misattributed to a co-actor than to a stranger – but only if the co-actor was present within the task environment.

The discovery of a robust self-prioritisation effect that is immune to the social context in which the task is performed, suggests that self-prioritisation may be a relatively 'hard-wired' aspect of human cognition. This theory is backed by neuroimaging and behavioural studies indicating that the self-relevance of information is determined relatively early in the processing stream, thus making it relatively resistant to modulation. For example, Sui et al. (2012) found that the degradation of visual stimuli, which is known to modulate relative early perceptual processing, has a weaker influence on processing if the stimulus is associated with the self than other individuals. Furthermore, the suppression of self-referent distractors activates the same regions of

the brain that are activated when supressing perceptually salient distractors (Sui, Rotshtein, et al., 2013). This was taken as evidence that self-relevance is processed in a similar (if not equivalent) manner as perceptual salience.

The finding that stimuli associated with a present co-actor were not prioritised for processing over stimuli associated with an imagined individual, suggests that the physical presence of the referent was not a factor used to prioritise social stimuli. This finding was counter to prediction on two fronts. First, I hypothesised that participants would prioritise stimuli associated with individuals who can effect immediate change in the environment. Second, I hypothesised that associations formed between a visual stimulus and a present individual would be more concrete than associations between a stimulus and an imagined individual. Psycholinguistic research has demonstrated that words representing concrete and imageable concepts are subject to a processing advantage relative to words representing abstract concepts (Fliessbach et al., 2006; Paivio et al., 1968). However, the present findings provide support that stimuli were prioritised based on relevance to the self only – and not on their physical distance or concreteness.

Although the physical presence of a co-actor did not influence match responses, there was an effect on non-match responses. Specifically, participants were slower to indicate that a self-associated stimulus had been paired with a non-matching label if that label referenced a present co-actor rather than to an absent stranger. This label effect was not found if both the stranger and the co-actor were absent from the task environment. I propose that this pattern of results reflects the degree of representational overlap between self and other. When attempting to categorise, or discriminate between, two or more items, conceptual similarity and perceptual similarity interact to influence decision making (Cheung & Gauthier, 2014; Humphreys & Forde, 2001; Wong & Wong, 2014). In perceptual matching tasks, the degree of perceptual similarity is controlled for, as shape and label assignment are counterbalanced across participants. However, the semantic classification of the stimuli (i.e., self versus other; present versus absent; self versus co-actor versus stranger) is the key independent variable. The more similar the semantic properties of the stimuli, the harder they are to discriminate

between. Therefore, the speed at which a participant determines that a shape label represents one of two different individuals will depend on how conceptually similar the two individuals are considered to be. The self and the present co-actor will have more representational similarity (i.e., both can be observed, both are generating salient events, both are considered participants, etc.) than the self and an imagined individual, whose representation with be lacking any specific detail. Therefore, discrimination judgements made between the concept of the self and a co-actor are more difficult if the co-actor is also directly observable.

In sum, the current research found no evidence that the self-prioritisation effect can be accounted for in terms of the physical presence of the self, relative to the physical absence of other individuals. The self-prioritisation effect remained just as robust in shared task environments as in solo task environments.

8.2 The role of 'likability' in self-prioritization

It has been proposed that self-bias could be equivalent to a positivity bias (de Greck et al., 2008; Enzi et al., 2009; Northoff & Hayes, 2011). That is, the intrinsic reward value or positive emotional valence of self-associated stimuli, biases cognitive processing. However, a participant's sensitivity to the self-relevance of stimuli was not found to correlate with their sensitivity to the reward value (Sui & Humphreys, 2015d) or affective valence (Stolte et al., 2017), of stimuli. This was taken as evidence that self-bias and positivity biases have distinct neural origins.

In Chapters 5 and 6, I used perceptual matching tasks to directly investigate whether stimuli associated with the self would be prioritised for processing over stimuli associated with individuals for whom the participant held either a favourable or unfavourable opinion. In Experiment 3, shape stimuli were associated with the self, a hero from popular culture (Harry Potter), and the affiliated villain (Lord Voldemort). In

Experiment 4 shape stimuli were associated with the self and two political adversaries (President Barack Obama and President Donald Trump). In Chapter 6 participants associated shape stimuli with the self, an unspecified stranger, and a self-selected politician for whom participants held a favourable (Experiment 5) or unfavourable (Experiment 6) opinion.

A robust self-prioritisation effect was found across experiments: participants were significantly faster to match shapes to the self than to either a liked or disliked other. Also, there was a consistent finding that participants were no faster to respond to stimuli associated with a liked-other than a disliked-other (or vice versa). Therefore, I found no evidence to suggest that visual stimuli in the perceptual matching task were prioritised based on the 'likability' of the person they represent. However, on trials where the self-associated shape was paired with a non-matching label, participants were significantly faster to perform a 'non-match' response when the self-associated shape referred to a disliked other than when it referred to a liked other.

I propose that participants are faster to disassociate themselves from disliked other than from a liked other due to a drive to maintain positive self-regard. It is a well-established phenomenon across social psychology that human behaviour is motivated by a need to establish and maintain self-esteem (Brown, 1986; Heine et al., 1999; James, 1890; Krueger, 1998). Furthermore, numerous studies have demonstrated that cognitive processes are biased to perceive oneself in a favourable light (Mezulis, Abramson, Hyde, & Hankin, 2004; Pahl & Eiser, 2005; Lynn A. Watson, Dritschel, Obonsawin, & Jentzsch, 2007). For example, people remember events in a way that emphasizes their most favourable attributes (Alicke & Govorun, 2005; Sedikides & Gregg, 2003, 2008). Therefore, within the context of the present study, it is likely that the cognitive system is distorted to emphasise differences between the self and a disliked-other, but similarities between the self and a liked-other.

In sum, the findings presented in Chapters 5 and 6 indicate that social stimuli are not prioritised for perceptual processing based on the 'likability' of

the referent. However, the perceptual and conceptual processing of social stimuli is distorted to distinguish the self from disliked others.

8.3 The prioritization of one's own task over the task of a present or absent co-actor.

In the final set of experiments, I investigated the extent to which participants prioritise their own task set over the task set of a co-actor. Pairs of participants were required to independently 'accept' or 'reject' a multi-feature stimulus based on whether it contained their individually assigned conjunction of target features. When participants were aware of their co-actor's task rules, the presence or absence of the co-actor's target features modulated their own-task performance. Specifically, when participants accepted the stimulus as containing their own target features, they were faster to do so if the co-actor's target features were also present in the stimulus. Conversely, when the participant was required to reject the stimulus as not containing their own target features, responses were significantly slower if the co-actor's target colour was present. When participants were unaware of their co-actor's task set, no such interference was found. The extent of cross-task interference was also significantly attenuated when the co-actor was absent from the task environment, but the participant still had an explicit awareness of the co-actor's task rules.

This experiment provided clear evidence that participants had not only represented the identity of their co-actor's target feature but had mapped it *specifically* to an accept response. Thus, when participants perceived the co-actor's target feature it primed a representation of an accept response within their own motor system. This priming facilitated responding when the participant was required by their own task set to accept the stimulus, but impaired responding when the participant was required to reject the stimulus. This is evidence that participants had formed an amalgamated 'co-representation' of the task, which incorporated their own task rules and their co-actor's task rules. An inability to fully distinguish their own task representation from their co-actor's task representation resulted in cross-task interference. Importantly, participants

had co-represented relatively complex binary-choice stimulus-response mappings, even though it provided no benefit to own-task performance. This pattern of results supports the idea that task co-representation is an unintended consequence of the social environment (Sebanz et al., 2005). That is, simply having an awareness of a nearby actor's task set automatically triggers task co-representation. Quasi-automatic task co-representation likely reflects the adaptive advantage of being able to anticipate the behaviors of nearby others – irrespective of costs to a person's own-task performance.

In sum, the experiments presented in Chapter 7 provide evidence that task co-representation extends well beyond simple stimulus-response mappings. Having an awareness of a co-actor's task rules modulates own task performance, irrespective of whether interpersonal coordination is required. Moreover, participants were not fully successful at prioritising the representation of their own task over the representation of the co-actor's task.

8.4 Overview

I now return to the main question of how self-prioritization research can inform us more generally about the mechanisms involved in social cognition. It has been well established that information associated with the self is subject to a substantial processing bias across cognitive domains. This finding has then leads to the further question of whether the self holds a unique status in cognition, or whether self-biases are driven by other general factors that are known to modulate processes of memory, perception and attention.

One such factor that may account for self-bias is familiarity. This is especially true of self-biases found in the processing of face stimuli, as performance may be modulated by the differential familiarity one has with their own face compared to the faces of known and unknown others (Keyes, Brady, Reilly, & Foxe, 2010; Sui, Hong, et

al., 2013; Sui et al., 2006). However, there is also evidence to suggest qualitative differences exist between the processing of one's own face and the faces of other individuals that do not reflect familiarity. For example, Keyes and Brady (2010) asked participants to classify face images of the self, a friend, or a stranger, that were either presented at upright (familiar) or inverted (unfamiliar) orientations. Participants were faster to respond to self-face images than other-face images, irrespective of whether the self-face image was presented upright or upside down. However, friend faces were recognized more quickly than unfamiliar faces, but only in upright, not inverted, orientations. Therefore, the results indicate that biases in self-face processing cannot fully be accounted for by stimulus familiarity.

Importantly, familiarity also cannot account for why people are faster to process novel stimuli that have only recently become associated with the self, relative to novel stimuli associated with familiar and unfamiliar others (Enock et al., 2018; Humphreys & Sui, 2015a; Sui et al., 2012). This finding has been reliably replicated across all the studies presented in chapters 4, 5 and 6. That is, people are markedly faster at responding to simple shape stimuli that have recently been associated with the self, then shape stimuli associated with familiar and unfamiliar others (co-actors, famous public figures, characters from popular fiction, and strangers). These studies have demonstrated a reliable self-association advantage even when the other people are highly familiar or even currently present in the task environment. Evidence across the perceptual matching studies presented within this thesis have indicated that, at the very least, there is a non-linear relationship between stimulus familiarity and self-bias.

Another possible explanation for why self-associated stimuli may be subject to enhanced processing is that associations formed between a stimulus and someone physically present in the environment might be more robust than associations with someone who is absent or hypothetical. That is, directly observing the self or others may reinforce and improve representations of that individual. This theory was tested in Chapter 4, in which participants associated novel shape stimuli with the self, present others, and absent others. There was no evidence to suggest the physical presence of an individual in the room strengthened associations between a novel stimulus and that

individual. While it is not possible to remove 'the self' from the environment, participants were no faster to respond to stimuli associated with a present other than an absent other. Therefore, it appears that self-bias is grounded within the mental representation of the self, and not just the physicality of the self.

It has also been proposed that self-prioritisation effects in social cognition may reflect the inherent reward value (Sui & Humphreys, 2015d) or emotional value (Stolte et al., 2017) of the self. If this is true, then it would be expected that self-associated stimuli would be processed in an equivalent manner to other high-reward or emotionally salient stimuli. However, while the self-association bias manifests in a similar manner to a reward or emotion bias, studies have indicated that they are not equivalent (Stolte et al., 2017; Sui & Humphreys, 2015d, 2015a). In particular, the magnitude of the self-biases demonstrated by an individual was not found to correlate with the magnitude of their biases towards the reward value or the emotional salience of stimuli.

Rather than taken this correlational approach between self-bias and a reward/emotion bias, experiments presented within Chapter 5 of this thesis directly investigated whether the magnitude of the self-association bias would differ if competing stimuli were associated with liked or disliked individuals. The size of the self-bias remained unaltered by the likability of the other individuals, indicating that an inherent positive association with the self, relative to other individuals, is not responsible for the self-prioritisation effect. This was further supported by the finding that participants were no faster to process information associated with liked-others over disliked-others. It is possible that the strength of the emotional association with the self, rather than the valence of the association, that is responsible for the enhanced processing of self-related information. However, the experiments presented in Chapter 6 revealed that participants were no faster to process stimuli associated with a liked/disliked other (a self-selected politician) than a neutral stranger. Overall, the evidence suggests that self-biases in social cognition cannot be accounted for by a generic positivity bias.

A recent proposed account for why self-reference has such a pervasive effect on cognition is that it acts as an 'associative glue', integrating perceptual and semantic information across processes of memory, attention, perception, and decision making (Sui & Humphreys, 2015c). This account would explain why shape and label stimuli associated with the self are matched faster than shape and label stimuli associated with other individuals (Sui et al., 2012).The idea of the self as an integrative glue is supported by memory studies showing that people demonstrate better recall for adjectives encoded in relation to the self, compared to adjectives encoded in relation to their semantic or physical properties (Bentley et al., 2017; Cunningham et al., 2014; Klein & Kihlstrom, 1986a; Klein & Loftus, 1988; Sui & Humphreys, 2015c). Furthermore, there is typically a higher degree of perceptual detail integrated in to memories encoded in reference to the self than memories encoded in reference to other individuals (Cunningham et al., 2014; Sui & Humphreys, 2015c; Sui & Zhu, 2005).

Neuroimaging studies have revealed that self-reference enhances the coupling between different stages of cognitive processing, such as the retrieval of a self-representation, and the allocation of attention to the environment (Sui & Humphreys, 2015c; Sui, Rotshtein, et al., 2013). A large-scale meta-analysis of neuroimaging studies, which contrasted neural correlates during processing of stimuli related to the self with those of non-self-referential stimuli, showed enhanced activation in the medial regions of the brains' cortex during the processing of self-related stimuli (Northoff et al., 2006). Importantly, activation of these cortical midline structures (CMS) occurred across functional domains (e.g., verbal, spatial, emotional, and facial). The CMS has also been associated with integrative processes involved in interoception (i.e., sensing the internal state of the body). Therefore, the perceptual distinction between self- and other-associated information may become bolstered by the internal sensations of having a body. On the basis of their meta-analysis, Northoff and colleagues concluded that self-referential processing in CMS likely constitutes the 'core of our self'. That is, the integration of self-related interoceptive (internal) and exteroceptive (external) stimuli results in a powerful sense of self that is distinct from others.

Why has the brain evolved to prioritise self-related information? When faced with 'perceptually noisy' environments containing more information than can possibly be processed, it is important to be able to selectively attend to and prioritise the processing of information most relevant to one's own goals. Information that (directly or indirectly) references the self is highly likely to be relevant to current goals. Therefore, the ability to quickly prioritise processing of self-referential information is advantageous. However, humans are also highly social creatures that often perform tasks in conjunction with other individuals. Therefore, it is not always beneficial to take an entirely egocentric view of the world. For instance, when attempting to carry a heavy object with another individual it is helpful not only to cognitively represent one's own share of the task, but also the co-actor's share of the task (Sebanz, Bekkering, et al., 2006b; Sebanz et al., 2003). In order to successfully understand the thoughts, beliefs, emotions and capabilities of another individual, it is necessary to temporarily supress one's self-representation and activate a representation of that individual's cognitive perspective.

Accordingly, the ability to not only identify with others, but also to distinguish between the self and other, is a crucial component of intersubjective transaction. There is a growing body of evidence from the fields of neuroscience, social psychology and developmental psychology, supporting the idea of a common representational network between self-representation and other-representation, both at the computational and neural levels (see Decety & Sommerville, 2003a for a review). The ability to concurrently hold two different representations relating to the self and others is known as co-representation (Wenke et al., 2011), and the ability to manipulate between the two representations is known as self-other control (de Guzman et al., 2016; Sowden & Shah, 2014).

An intriguing finding that questions the success of self-other control is that people demonstrate interference from a nearby co-actor's task even when no interpersonal coordination is required or expected. That is, even in complete division of labour scenarios in which focusing entirely on one's own task component would result in optimal task performance, having an awareness of a co-actor's task set results in

cross-task interference (e.g., Atmaca et al., 2011; Böckler et al., 2012; Sebanz et al., 2003). Therefore, even when it would be beneficial to prioritise information relating to one's own task over a representation of another individual's task, self-prioritisation is not fully achieved.

It has been proposed that task co-representation is an automatic and socially-driven aspect of human cognition (Knoblich & Sebanz, 2006; Sebanz et al., 2003, 2005). That is, people performing tasks within a shared environment automatically represent the actions and task sets of those around them. However, the automaticity of task co-representation has been called in to question due to limitations of the paradigms used to investigate shared-task performance (e.g., Dolk et al., 2011, 2013). For example, the most common task used to investigate task co-representation, the 'Social' or 'Joint' Simon task (Sebanz et al., 2003), requires co-actors to take turns to respond to different elements of the same stimulus. Due to the complementary nature of the responses, knowing when a co-actor should perform an action is equivalent to knowing when to withhold from responding. Therefore, representing a co-actor's task component may be a deliberate, and not an automatic, strategy. Furthermore, studies adopting the social simon task typically report greater interference when co-actors perform their complementary task components together, then when they perform them apart (see Dolk, Hommel, Colzato, et al., 2014 for a review). Introducing the co-actor to the task environment causes fundamental changes to the task environment that may account for the differences in interference effects (Dittrich, Rothe, & Klauer, 2012; Dolk et al., 2013).

The experiments presented within chapter 7 used a novel paradigm that removed confounds relating to turn taking or changes to the task environment. They provide convincing evidence that people do, in fact, take in to account a co-actor's task rules when there is no benefit of doing so. Furthermore, participants were found to co-represent complex and indirect stimulus-response mappings, resulting in the modulation of own-task performance. How do findings of pervasive and sizeable self-prioritisation effects across cognitive domains fit with this finding that people are unable to completely prioritise their own task instruction when in shared environments? The

answer likely relates to the activation of a 'self-other' control mechanism (de Guzman et al., 2016; Sowden & Shah, 2014).

The ability to hold in mind two concurrent mental representations relating to the self and other, and carefully balance activation of the two, is thought to recruit a specific control mechanism (Sowden & Shah, 2014). This mechanism is necessary to relate the self to others (i.e., during empathy, theory of mind, and joint action) and to distinguish the self from others (i.e., during self-prioritisation). Difficulties with self-other control have been associated with a number of socio-cognitive disorders, such as autism spectrum disorder (ASD), schizophrenia, and mirror-touch synaesthesia. For instance, some of the characteristics of ASD include difficulties in attributing mental states to others, and the involuntary repetitions of observed actions and speech (Baron-Cohen, 1997; Begeer, Bernstein, van Wijhe, Scheeren, & Koot, 2012; Kana et al., 2017; Prizant & Duchan, 1981).

Even within neurotypical populations, there is evidence that a form of 'social contagion' exists. That is, people demonstrate an automatic tendency to imitate others when in shared environments (Heyes, 2011; Iacoboni, 2009). It has been demonstrated that people are faster to initiate an action if they observe someone performing the same action, than if they observe someone performing a different action (Brass, Derrfuss, & von Cramon, 2005; Brass et al., 2009). It is important to note that healthy individuals do not get stuck in imitative loops indefinitely. Therefore, top-down control is exerted to supress imitation in order to execute one's own action plan. This is an example of prioritising a representation of one's own action intentions over a representation of an observed action (self-prioritisation). In contrast, when performing a joint action with another individual, representations of the co-actor's action plan must be activated in order to plan and coordinate behaviours (Sebanz, Bekkering, et al., 2006b; Wenke et al., 2011). This ability to carefully relate and manipulate representations of self and other depending on task demands is evidence of a self-other control mechanism.

Overall, evidence suggests that there are two important drives in social cognition: (1) the drive to prioritise information most relevant to one's own goals, and

(2) the drive to understand and represent the thoughts, emotions and action plans of others. In certain contexts it is relevant to one's own goals to be aware of information that can inform us of what nearby actors are doing. Therefore, there is evidence of a flexible mechanism that can control the prioritisation of self- and other- related information across domains of social cognition.

8.5 Concluding remarks

The research presented in this thesis extends and provides new knowledge of the role of the self in the processing and organizing of incoming information. It is clear that the self-relevance of stimuli is a powerful determinant of prioritized perceptual processing. However, within the context of co-action and task sharing, there is also a prevailing drive to anticipate the actions of nearby others. Therefore, an awareness of how a stimulus relates to the actions of a nearby individual can influence one's own action plans, irrespective of whether interpersonal coordination is warranted.

8.6 Potential future directions and remaining questions

Although a lot has been learned, the area of self-prioritisation remains an important and exciting topic with much still left to discover. To conclude this thesis, below I list three from a vast range of possible remaining questions for future exploration.

- Being 'a self' is both a universal and an idiosyncratic experience. Individual differences in self-representation, and how they relate to self-prioritization effects, remains a topic to be explored. For example, is the strength of self-bias correlated with traits such as empathy or introversion?

- At what point do hypo and hyper self-biases result in socio-cognitive impairments?
- There are also interesting questions relating to the adoption of other 'selves' in virtual reality environments. In virtual reality it is now possible to experience an environment through an avatar that can vary dramatically from one's real-world self in terms of names, appearance, capabilities, and even personality traits. To what extent does one's self concept adapt to a virtual self?

9 REFERENCES

Aboulafia-Brakha, T., Christe, B., Martory, M.-D., & Annoni, J.-M. (2011). Theory of mind tasks and executive functions: a systematic review of group studies in neurology. *Journal of Neuropsychology*, *5*(Pt 1), 39–55. https://doi.org/10.1348/174866410X533660

Abu-Akel, A., & Shamay-Tsoory, S. G. (2013). Characteristics of theory of mind impairments in schizophrenia. *Social Cognition in Schizophrenia: From Evidence to Treatment*, 196–214.

Alexopoulos, T., Muller, D., Ric, F., & Marendaz, C. (2012). I, me, mine: Automatic attentional capture by self-related stimuli: Attention to self-related stimuli. *European Journal of Social Psychology*, *42*(6), 770–779. https://doi.org/10.1002/ejsp.1882

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The Self in Social Judgment*, *1*, 85–106.

Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, *20*(1), 1–48. https://doi.org/10.1080/10463280802613866

Allan, K., Morson, S., Dixon, S., Martin, D., & Cunningham, S. J. (2017). Simulation-based mentalizing generates a "proxy" self-reference effect in memory. *The Quarterly Journal of Experimental Psychology*, *70*(6), 1074–1084.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.

Anderson, A. K. (2005). Affective influences on the attentional dynamics supporting awareness. *Journal of Experimental Psychology. General*, *134*(2), 258–281. https://doi.org/10.1037/0096-3445.134.2.258

Anderson, B. A., & Yantis, S. (2013). Persistence of value-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 6.

Apperly, I. A. (2010). *Mindreaders: the cognitive basis of" theory of mind"*. Psychology Press.

Apperly, I. A., Samson, D., Chiavarino, C., Bickerton, W.-L., & Humphreys, G. W. (2007). Testing the domain-specificity of a theory of mind deficit in brain-injured patients: Evidence for consistent performance on non-verbal,"reality-unknown" false belief and false photograph tasks. *Cognition*, *103*(2), 300–321.

Arnell, K. M., Shapiro, K. L., & Sorensen, R. E. (1999). Reduced Repetition Blindness for One's Own Name. *Visual Cognition*, *6*(6), 609–635. https://doi.org/10.1080/135062899394876

Atmaca, S., Sebanz, N., & Knoblich, G. (2011). The joint flanker effect: sharing tasks with real and imagined co-actors. *Experimental Brain Research*, *211*(3–4), 371–385. https://doi.org/10.1007/s00221-011-2709-9

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: a 1-year longitudinal perspective. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00655

Avenanti, A., Bueti, D., Galati, G., & Aglioti, S. M. (2005). Transcranial magnetic stimulation highlights the sensorimotor side of empathy for pain. *Nature Neuroscience*, *8*(7), 955.

Banissy, M. J., Kadosh, R. C., Maus, G. W., Walsh, V., & Ward, J. (2009). Prevalence, characteristics and a neurocognitive model of mirror-touch synaesthesia. *Experimental Brain Research*, *198*(2–3), 261–272.

Banissy, M. J., & Ward, J. (2007). Mirror-touch synesthesia is linked with empathy. *Nature Neuroscience*, *10*(7), 815.

Banissy, M. J., & Ward, J. (2013). Mechanisms of self-other representations and vicarious experiences of touch in mirror-touch synesthesia. *Frontiers in Human Neuroscience*, *7*. https://doi.org/10.3389/fnhum.2013.00112

Baron, R. S. (1986). Distraction-conflict theory: Progress and problems. In *Advances in experimental social psychology* (Vol. 19, pp. 1–40). Elsevier.

Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*(1), 37–46.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163–175.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition*, *57*(3), 217–239. https://doi.org/10.1016/0010-0277(95)00669-X

Begeer, S., Bernstein, D. M., van Wijhe, J., Scheeren, A. M., & Koot, H. M. (2012). A continuous false belief task reveals egocentric biases in children and adolescents with Autism Spectrum Disorders. *Autism*, *16*(4), 357–366.

Beggan, J. K. (1992). On the social nature of nonsocial perception: The mere ownership effect. *Journal of Personality and Social Psychology*, *62*(2), 229–237. https://doi.org/10.1037/0022-3514.62.2.229

Bentley, S. V., Greenaway, K. H., & Haslam, S. A. (2017). An online paradigm for exploring the self-reference effect. *PLOS ONE*, *12*(5), e0176611. https://doi.org/10.1371/journal.pone.0176611

Biedermann, F., Frajo-Apor, B., & Hofer, A. (2012). Theory of mind and its relevance in schizophrenia. *Current Opinion in Psychiatry*, *25*(2), 71–75.

Blaine, B., & Crocker, J. (1993). Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review. In *Self-esteem* (pp. 55–85). Springer.

Blakemore, S.-J., Bristow, D., Bird, G., Frith, C. D., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. *Brain*, *128*(7), 1571–1583.

Blakemore, S.-J., & Frith, C. D. (2003). Self-awareness and action. *Current Opinion in Neurobiology*, *13*(2), 219–224.

Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 582.

Böckler, A., Knoblich, G., & Sebanz, N. (2012). Effects of a coactor's focus of attention on task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(6), 1404–1415. https://doi.org/10.1037/a0027523

Bora, E., & Pantelis, C. (2013). Theory of mind impairments in first-episode psychosis, individuals at ultra-high risk for psychosis and in first-degree relatives of schizophrenia: systematic review and meta-analysis. *Schizophrenia Research*, *144*(1), 31–36.

Borroni, P., & Baldissera, F. (2008). Activation of motor pathways during observation and execution of hand movements. *Social Neuroscience*, *3*(3–4), 276–288. https://doi.org/10.1080/17470910701515269

Bradford, E. E., Brunsdon, V. E., & Ferguson, H. J. (2017). The Relationship Between Theory of Mind and Executive Functions in Older and Younger Adults.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between Observed and Executed Finger Movements: Comparing Symbolic, Spatial, and Imitative Cues. *Brain and Cognition*, *44*(2), 124–143. https://doi.org/10.1006/brcg.2000.1225

Brass, M., Derrfuss, J., Matthes-von Cramon, G., & von Cramon, D. Y. (2003). Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology*, *17*(2), 265–271. https://doi.org/10.1037/0894-4105.17.2.265

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative and overlearned responses: a functional double dissociation. *Neuropsychologia*, *43*(1), 89–98. https://doi.org/10.1016/j.neuropsychologia.2004.06.018

Brass, M., & Heyes, C. (2005). Imitation: is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences*, *9*(10), 489–495.

Brass, M., Ruby, P., & Spengler, S. (2009). Inhibition of imitative behaviour and social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2359–2367. https://doi.org/10.1098/rstb.2009.0066

Broadbent, D. E. (1958). *Perception and communication.* Elmsford, NY, US: Pergamon Press, Inc. https://doi.org/10.1037/10037-000

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition*, *4*(4), 353.

Campoy, G., Castellà, J., Provencio, V., Hitch, G. J., & Baddeley, A. D. (2015). Automatic semantic encoding in verbal short-term memory: Evidence from the

concreteness effect. *The Quarterly Journal of Experimental Psychology*, *68*(4), 759–778. https://doi.org/10.1080/17470218.2014.966248

Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The Relation Between Executive Function and Theory of Mind is More Than Skin Deep. *Journal of Cognition and Development*, *16*(1), 186–197. https://doi.org/10.1080/15248372.2013.824883

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, *72*(4), 1032–1053.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development*, *11*(2), 73–92.

Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*(4), 299–319. https://doi.org/10.1016/j.jecp.2004.01.002

Castelli, F., Frith, C. D., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, *125*(8), 1839–1849.

Catmur, C., Walsh, V., & Heyes, C. (2009). Associative sequence learning: the role of experience in the development of imitation and the mirror system. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1528), 2369–2380. https://doi.org/10.1098/rstb.2009.0048

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979.

Cheung, O. S., & Gauthier, I. (2014). Visual appearance interacts with conceptual knowledge in object recognition. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00793

Cho, R., & Wu, W. (2014). Is inner speech the basis of auditory verbal hallucination in schizophrenia? *Frontiers in Psychiatry*, *5*, 75.

Constable, M. D., Pratt, J., Gozli, D. G., & Welsh, T. N. (2015). Do you see what I see? Co-actor posture modulates visual processing in joint tasks. *Visual Cognition*, *23*(6), 699–719. https://doi.org/10.1080/13506285.2015.1078426

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*(2), 331–335.

Conway, M. A. (2005). Memory and the self. *Journal of Memory and Language*, *53*(4), 594–628. https://doi.org/10.1016/j.jml.2005.08.005

Cook, J., Barbalat, G., & Blakemore, S.-J. (2012). Top-down modulation of the perception of other people in schizophrenia and autism. *Frontiers in Human Neuroscience*, *6*, 175.

Cook, J., & Bird, G. (2012). Atypical social modulation of imitation in autism spectrum conditions. *Journal of Autism and Developmental Disorders*, *42*(6), 1045–1051. https://doi.org/10.1007/s10803-011-1341-7

Cook, R., Bird, G., Lünser, G., Huck, S., & Heyes, C. (2011). Automatic imitation in a strategic context: players of rock–paper–scissors imitate opponents' gestures. *Proceedings of the Royal Society of London B: Biological Sciences*, rspb20111024.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684. https://doi.org/10.1016/S0022-5371(72)80001-X

Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268.

Cross, K. A., & Iacoboni, M. (2014). Neural systems for preparatory control of imitation. *Phil. Trans. R. Soc. B*, *369*(1644), 20130176. https://doi.org/10.1098/rstb.2013.0176

Cross, K. A., Torrisi, S., Reynolds Losin, E. A., & Iacoboni, M. (2013). Controlling automatic imitative tendencies: Interactions between mirror neuron and cognitive control systems. *NeuroImage*, *83*, 493–504. https://doi.org/10.1016/j.neuroimage.2013.06.060

Cunningham, S. J., Brebner, J. L., Quinn, F., & Turk, D. J. (2014). The self-reference effect on memory in early childhood. *Child Development*, *85*(2), 808–823.

Cunningham, S. J., & Turk, D. J. (2017). A review of self-processing biases in cognition. Retrieved from https://repository.abertay.ac.uk/jspui/handle/10373/2580

Czienskowski, U., & Giljohann, S. (2002). Intimacy, concreteness, and the 'self-reference effect'. *Experimental Psychology*, *49*(1), 73–79. https://doi.org/10.1027//1618-3169.49.1.73

Damon, W., & Hart, D. (1982). The development of self-understanding from infancy through adolescence. *Child Development*, 841–864.

Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, *9*(1), 28.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078. https://doi.org/10.1016/j.neuropsychologia.2006.02.006

de Caso, I., Poerio, G., Jefferies, E., & Smallwood, J. (2017). That's me in the spotlight: neural basis of individual differences in self-consciousness. *Social Cognitive and Affective Neuroscience*, *12*(9), 1384–1393. https://doi.org/10.1093/scan/nsx076

de Greck, M., Rotte, M., Paus, R., Moritz, D., Thiemann, R., Proesch, U., … Northoff, G. (2008). Is our self based on reward? Self-relatedness recruits neural activity in the reward system. *NeuroImage*, *39*(4), 2066–2075. https://doi.org/10.1016/j.neuroimage.2007.11.006

de Guzman, M., Bird, G., Banissy, M. J., & Catmur, C. (2016). Self–other control processes in social cognition: from imitation to empathy. *Phil. Trans. R. Soc. B*, *371*(1686), 20150079.

DeBruine, L. M., Jones, B. C., Little, A. C., & Perrett, D. I. (2008). Social perception of facial resemblance in humans. *Archives of Sexual Behavior*, *37*(1), 64–77.

Decety, J., & Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, *3*(2), 71–100.

Decety, J., & Lamm, C. (2007). The Role of the Right Temporoparietal Junction in Social Interaction: How Low-Level Computational Processes Contribute to Meta-Cognition. *The Neuroscientist*, *13*(6), 580–593. https://doi.org/10.1177/1073858407304654

Decety, J., & Lamm, C. (2009). *Empathy versus Personal Distress: Recent Evidence from Social Neuroscience*. https://doi.org/10.7551/mitpress/9780262012973.003.0016

Decety, J., & Sommerville, J. A. (2003a). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527–533. https://doi.org/10.1016/j.tics.2003.10.004

Decety, J., & Sommerville, J. A. (2003b). Shared representations between self and other: a social cognitive neuroscience view. *Trends in Cognitive Sciences*, *7*(12), 527–533. https://doi.org/10.1016/j.tics.2003.10.004

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, *41*(5), 665–678.

Dittrich, K., Rothe, A., & Klauer, K. C. (2012). Increased spatial salience in the social Simon task: A response-coding account of spatial compatibility effects. *Attention, Perception, & Psychophysics*, *74*(5), 911–929. https://doi.org/10.3758/s13414-012-0304-1

Dolcos, F., LaBar, K. S., & Cabeza, R. (2004). Dissociable effects of arousal and valence on prefrontal activity indexing emotional evaluation and subsequent memory: an event-related fMRI study. *NeuroImage*, *23*(1), 64–74. https://doi.org/10.1016/j.neuroimage.2004.05.015

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2011). How "Social" is the social Simon effect? *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00084

Dolk, T., Hommel, B., Colzato, L. S., Schütz-Bosbach, S., Prinz, W., & Liepelt, R. (2014). The joint Simon effect: a review and theoretical integration. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00974

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2013). The (not so) social Simon effect: A referential coding account. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1248–1260. https://doi.org/10.1037/a0031031

Dolk, T., Hommel, B., Prinz, W., & Liepelt, R. (2014). The joint flanker effect: Less social than previously thought. *Psychonomic Bulletin & Review*, *21*(5), 1224–1230. https://doi.org/10.3758/s13423-014-0583-8

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Attention, Perception, & Psychophysics*, *63*(6), 1004–1013.

Eisenberg, N., & Eggum, N. D. (2009). Empathic responding: Sympathy and personal distress. *The Social Neuroscience of Empathy*, *6*, 71–83.

Enock, F., Sui, J., Hewstone, M., & Humphreys, G. W. (2018). Self and team prioritisation effects in perceptual matching: Evidence for a shared representation. *Acta Psychologica*, *182*, 107–118. https://doi.org/10.1016/j.actpsy.2017.11.011

Enzi, B., de Greck, M., Prösch, U., Tempelmann, C., & Northoff, G. (2009). Is our self nothing but reward? Neuronal overlap and distinction between reward and personal relevance and its relation to human personality. *PloS One*, *4*(12), e8429. https://doi.org/10.1371/journal.pone.0008429

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327.

Etcoff, N. L., & Magee, J. J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227–240. https://doi.org/10.1016/0010-0277(92)90002-Y

Eysenck, M. W., & Eysenck, M. C. (1979). Processing depth, elaboration of encoding, memory stores, and expended processing capacity. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(5), 472.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, *73*(6), 2608–2611.

Fan, Y., Duncan, N. W., de Greck, M., & Northoff, G. (2011). Is there a core neural network in empathy? An fMRI based quantitative meta-analysis. *Neuroscience & Biobehavioral Reviews*, *35*(3), 903–911.

Fan, Y.-T., Decety, J., Yang, C.-Y., Liu, J.-L., & Cheng, Y. (2010). Unbroken mirror neurons in autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *51*(9), 981–988.

Farrer, C., Franck, N., Frith, C. D., Decety, J., Georgieff, N., d'Amato, T., & Jeannerod, M. (2004). Neural correlates of action attribution in schizophrenia. *Psychiatry Research*, *131*(1), 31–44. https://doi.org/10.1016/j.pscychresns.2004.02.004

Fassbender, C., Simoes-Franklin, C., Murphy, K., Hester, R., Meaney, J., Robertson, I. H., & Garavan, H. (2006). The role of a right fronto-parietal network in cognitive control. *Journal of Psychophysiology*, *20*(4), 286–296.

Feinberg, J. M., & Aiello, J. R. (2006). Social facilitation: A test of competing theories. *Journal of Applied Social Psychology*, *36*(5), 1087–1109.

Fields, E. C., & Kuperberg, G. R. (2015). Loving yourself more than your neighbor: ERPs reveal online effects of a self-positivity bias. *Social Cognitive and Affective Neuroscience*, *10*(9), 1202–1209. https://doi.org/10.1093/scan/nsv004

Finke, J. B., Larra, M. F., Merz, M. U., & Schächinger, H. (2017). Startling similarity: Effects of facial self-resemblance and familiarity on the processing of emotional faces. *PLOS ONE*, *12*(12), e0189028. https://doi.org/10.1371/journal.pone.0189028

Fitzgibbon, B. M., Giummarra, M. J., Georgiou-Karistianis, N., Enticott, P. G., & Bradshaw, J. L. (2010). Shared pain: from empathy to synaesthesia. *Neuroscience & Biobehavioral Reviews*, *34*(4), 500–512.

Fizke, E., Barthel, D., Peters, T., & Rakoczy, H. (2014). Executive function plays a role in coordinating different perspectives, particularly when one's own perspective is involved. *Cognition*, *130*(3), 315–334.

Fliessbach, K., Weis, S., Klaver, P., Elger, C. E., & Weber, B. (2006). The effect of word concreteness on recognition memory. *NeuroImage*, *32*(3), 1413–1421.

Fonagy, P., & Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology*, *21*(4), 1355–1381.

Forsyth, D. R. (2008). Self-serving bias.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial Expressions of Emotion: Are Angry Faces Detected More Efficiently? *Cognition & Emotion*, *14*(1), 61–92. https://doi.org/10.1080/026999300378996

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex*, *86*, 186–204. https://doi.org/10.1016/j.cortex.2016.04.023

Frith, C. D. (2014). *The cognitive neuropsychology of schizophrenia*. Psychology press.

Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Phil. Trans. R. Soc. Lond. B*, *355*(1404), 1771–1788.

Frith, C. D., & Frith, U. (2005). Theory of mind. *Current Biology*, *15*(17), R644–R645.

Frith, C. D., & Frith, U. (2006a). How we predict what other people are going to do. *Brain Research*, *1079*(1), 36–46.

Frith, C. D., & Frith, U. (2006b). The Neural Basis of Mentalizing. *Neuron*, *50*(4), 531–534. https://doi.org/10.1016/j.neuron.2006.05.001

Frith, C. D., & Frith, U. (2007). Social Cognition in Humans. *Current Biology*, *17*(16), R724–R732. https://doi.org/10.1016/j.cub.2007.05.068

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 459–473.

Galinsky, A. D., & Mussweiler, T. (2001). First offers as anchors: the role of perspective-taking and negotiator focus. *Journal of Personality and Social Psychology*, *81*(4), 657.

Gallese, V. (2001). The'shared manifold'hypothesis. From mirror neurons to empathy. *Journal of Consciousness Studies*, *8*(5–6), 33–50.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*(12), 493–501.

Gangitano, M., Mottaghy, F. M., & Pascual-Leone, A. (2001). Phase-specific modulation of cortical motor output during movement observation: *Neuroreport*, *12*(7), 1489–1492. https://doi.org/10.1097/00001756-200105250-00038

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11.

Gaspelin, N., Ruthruff, E., & Jung, K. (2014). Slippage theory and the flanker paradigm: An early-selection account of selective attention failures. *Journal of Experimental Psychology: Human Perception and Performance*, *40*(3), 1257.

Geng, H., Zhang, S., Li, Q., Tao, R., & Xu, S. (2012). Dissociations of subliminal and supraliminal self-face from other-face processing: Behavioral and ERP evidence. *Neuropsychologia*, *50*(12), 2933–2942. https://doi.org/10.1016/j.neuropsychologia.2012.07.040

Genschow, O., Bossche, S. van D., Cracco, E., Bardi, L., Rigoni, D., & Brass, M. (2017). Mimicry and automatic imitation are not correlated. *PLOS ONE*, *12*(9), e0183784. https://doi.org/10.1371/journal.pone.0183784

Giesbrecht, B., Sy, J. L., & Lewis, M. K. (2009). Personal names do not always survive the attentional blink: Behavioral evidence for a flexible locus of selection. *Vision Research*, *49*(10), 1378–1388. https://doi.org/10.1016/j.visres.2008.02.013

Goldstone, R. L., & Hendrickson, A. T. (2010a). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69–78.

Goldstone, R. L., & Hendrickson, A. T. (2010b). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(1), 69–78.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001a). Altering object representations through category learning. *Cognition*, *78*(1), 27–43.

Goldstone, R. L., Lippa, Y., & Shiffrin, R. M. (2001b). Altering object representations through category learning. *Cognition*, *78*(1), 27–43.

Gopnik, A., & Astington, J. W. (1988). Children's Understanding of Representational Change and Its Relation to the Understanding of False Belief and the Appearance-Reality Distinction. *Child Development*, *59*(1), 26–37. https://doi.org/10.2307/1130386

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological Review*, *102*(1), 4.

Gronau, N., Cohen, A., & Ben-Shakhar, G. (2003). Dissociations of Personally Significant and Task-Relevant Distractors Inside and Outside the Focus of Attention: A Combined Behavioral and Psychophysiological Study. *Journal of*

Experimental Psychology: General, *132*(4), 512–529. https://doi.org/10.1037/0096-3445.132.4.512

Gudykunst, W. B., Matsumoto, Y., Ting-Toomey, S., Nishida, T., Kim, K., & Heyman, S. (1996). The influence of cultural individualism-collectivism, self construals, and individual values on communication styles across cultures. *Human Communication Research*, *22*(4), 510–543.

Hala, S., Hug, S., & Henderson, A. (2003). Executive function and false-belief understanding in preschool children: Two tasks are harder than one. *Journal of Cognition and Development*, *4*(3), 275–298.

Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: what can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain*, *4*(2), 159. https://doi.org/10.1007/s40167-016-0044-5

Hamilton, M., & Rajaram, S. (2001a). The concreteness effect in implicit and explicit memory tests. *Journal of Memory and Language*, *44*(1), 96–117.

Hamilton, M., & Rajaram, S. (2001b). The Concreteness Effect in Implicit and Explicit Memory Tests. *Journal of Memory and Language*, *44*(1), 96–117. https://doi.org/10.1006/jmla.2000.2749

Han, S., & Northoff, G. (2009). Understanding the self: a cultural neuroscience approach. In J. Y. Chiao (Ed.), *Progress in Brain Research* (Vol. 178, pp. 203–212). Elsevier. https://doi.org/10.1016/S0079-6123(09)17814-7

Hari, R., Himberg, T., Nummenmaa, L., Hämäläinen, M., & Parkkonen, L. (2013). Synchrony of brains and bodies during implicit interpersonal interaction. *Trends in Cognitive Sciences*, *17*(3), 105–106.

Harris, C. R., Pashler, H. E., & Coburn, P. (2004). Moray revisited: High-priority affective stimuli and visual search. *The Quarterly Journal of Experimental Psychology Section A*, *57*(1), 1–31. https://doi.org/10.1080/02724980343000107

Heine, S. J., Lehman, D. R., Markus, H. R., & Kitayama, S. (1999). Is there a universal need for positive self-regard? *Psychological Review*, *106*(4), 766–794.

Heyes, C. (2011). Automatic imitation. *Psychological Bulletin*, *137*(3), 463–483. https://doi.org/10.1037/a0022288

Hickey, C., Chelazzi, L., & Theeuwes, J. (2011). Reward has a residual impact on target selection in visual search, but not on the suppression of distractors. *Visual Cognition*, *19*(1), 117–128.

Hoffmann, F., Banzhaf, C., Kanske, P., Gärtner, M., Bermpohl, F., & Singer, T. (2016). Empathy in depression: Egocentric and altercentric biases and the role of alexithymia. *Journal of Affective Disorders*, *199*, 23–29. https://doi.org/10.1016/j.jad.2016.03.007

Hofree, G., Urgen, B. A., Winkielman, P., & Saygin, A. P. (2015). Observation and imitation of actions performed by humans, androids, and robots: an EMG study. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00364

Holle, H., Banissy, M., Wright, T., Bowling, N., & Ward, J. (2011). 'That's not a real body': identifying stimulus qualities that modulate synaesthetic experiences of touch. *Consciousness and Cognition*, *20*(3), 720–726. https://doi.org/10.1016/j.concog.2010.12.002

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 65–70.

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research*, *59*(3), 176–186.

Hommel, B. (2011). The Simon effect as tool and heuristic. *Acta Psychologica*, *136*(2), 189–202. https://doi.org/10.1016/j.actpsy.2010.04.011

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). Codes and their vicissitudes. *Behavioral and Brain Sciences*, *24*(5), 910–926.

Huang, C.-M., Andrist, S., Sauppé, A., & Mutlu, B. (2015). Using gaze patterns to predict task intent in collaboration. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.01049

Humphreys, G. W., & Forde, E. M. (2001). Hierarchies, similarity, and interactivity in object recognition: 'category-specific' neuropsychological deficits. *The Behavioral and Brain Sciences*, *24*(3), 453–476; discussion 476-509.

Humphreys, G. W., & Sui, J. (2015a). The salient self: Social saliency effects based on self-bias. *Journal of Cognitive Psychology*, *27*(2), 129–140. https://doi.org/10.1080/20445911.2014.996156

Humphreys, G. W., & Sui, J. (2015b). The salient self: Social saliency effects based on self-bias. *Journal of Cognitive Psychology*, *27*(2), 129–140. https://doi.org/10.1080/20445911.2014.996156

Humphreys, G. W., & Sui, J. (2016). Attentional control and the self: The Self-Attention Network (SAN). *Cognitive Neuroscience*, *7*(1–4), 5–17. https://doi.org/10.1080/17588928.2015.1044427

Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual Review of Psychology*, *60*, 653–670.

Iacoboni, M., & Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience*, *7*(12), 942.

Igelström, K. M., Webb, T. W., & Graziano, M. S. A. (2017). Functional Connectivity Between the Temporoparietal Cortex and Cerebellum in Autism Spectrum Disorder. *Cerebral Cortex (New York, N.Y.: 1991)*, *27*(4), 2617–2627. https://doi.org/10.1093/cercor/bhw079

Itier, R. J., & Taylor, M. J. (2004). N170 or N1? Spatiotemporal Differences between Object and Face Processing Using ERPs. *Cerebral Cortex*, *14*(2), 132–142. https://doi.org/10.1093/cercor/bhg111

Jackson, P. L., Rainville, P., & Decety, J. (2006). To what extent do we share the pain of others? Insight from the neural bases of pain empathy. *Pain*, *125*(1), 5–9.

James, W. (1890). *The Principles of Psychology*. Dover Publications.

James, W. (1950). *The principles of psychology (Vols. 1 & 2)*. New York: Dover. (Original work published 1890).

Jardri, R., Pins, D., Lafargue, G., Very, E., Ameller, A., Delmaire, C., & Thomas, P. (2011). Increased overlap between the brain areas involved in self-other distinction in schizophrenia. *PloS One*, *6*(3), e17500.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, *98*(6), 1325–1348. https://doi.org/10.1086/261737

Kana, R. K., Sartin, E. B., Stevens Jr, C., Deshpande, H. D., Klein, C., Klinger, M. R., & Klinger, L. G. (2017). Neural networks underlying language and social cognition during self-other processing in Autism spectrum disorders. *Neuropsychologia*, *102*, 116–123.

Kanske, P., Böckler, A., Trautwein, F.-M., Parianen Lesemann, F. H., & Singer, T. (2016). Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive and Affective Neuroscience*, *11*(9), 1383–1392. https://doi.org/10.1093/scan/nsw052

Kanske, P., Böckler, A., Trautwein, F.-M., & Singer, T. (2015). Dissecting the social brain: Introducing the EmpaToM to reveal distinct neural networks and brain-behavior relations for empathy and Theory of Mind. *NeuroImage*, *122*, 6–19. https://doi.org/10.1016/j.neuroimage.2015.07.082

Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*(5), 785–794. https://doi.org/10.1162/08989290260138672

Keyes, H., & Brady, N. (2010). Self-face recognition is characterized by "bilateral gain" and by faster, more accurate performance which persists when faces are inverted. *Quarterly Journal of Experimental Psychology*, *63*(5), 840–847. https://doi.org/10.1080/17470211003611264

Keyes, H., Brady, N., Reilly, R. B., & Foxe, J. J. (2010). My face or yours? Event-related potential correlates of self-face processing. *Brain and Cognition*, *72*(2), 244–254. https://doi.org/10.1016/j.bandc.2009.09.006

Keyes, H., & Dlugokencka, A. (2014). Do I Have My Attention? Speed of Processing Advantages for the Self-Face Are Not Driven by Automatic Attention Capture. *PLOS ONE*, *9*(10), e110792. https://doi.org/10.1371/journal.pone.0110792

Kiernan, D. S., Neyedli, H. F., Ray, M., Potruff, A., Pratt, J., Weeks, D. J., & Welsh, T. N. (2012). 'You were always on my mind': Action co-representation in Joint Simon tasks. *Journal of Vision*, *12*(9), 821–821. https://doi.org/10.1167/12.9.821

Kilner, J. M., Vargas, C., Duval, S., Blakemore, S.-J., & Sirigu, A. (2004). Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, *7*(12), 1299. https://doi.org/10.1038/nn1355

Kirita, T., & Endo, M. (1995). Happy face advantage in recognizing facial expressions. *Acta Psychologica*, *89*(2), 149–163. https://doi.org/10.1016/0001-6918(94)00021-8

Kirova, A.-M., Bays, R. B., & Lagalwar, S. (2015). Working Memory and Executive Function Decline across Normal Aging, Mild Cognitive Impairment, and Alzheimer's Disease. *BioMed Research International*, *2015*, 1–9. https://doi.org/10.1155/2015/748212

Klein, & Kihlstrom. (1986a). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology. General*, *115*(1), 26–38.

Klein, S. B. (2012). Self, memory, and the self-reference effect: an examination of conceptual and methodological issues. *Personality and Social Psychology Review*, *16*(3), 283–300.

Klein, S. B., & Kihlstrom, J. F. (1986b). Elaboration, organization, and the self-reference effect in memory. *Journal of Experimental Psychology: General*, *115*(1), 26–38. https://doi.org/10.1037//0096-3445.115.1.26

Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology*, *55*(1), 5–11. https://doi.org/10.1037/0022-3514.55.1.5

Klimecki, O. M., Vuilleumier, P., & Sander, D. (2016). The impact of emotions and empathy-related traits on punishment behavior: introduction and validation of the inequality game. *PLoS One*, *11*(3), e0151028.

Knetsch, J. L. (1989). The Endowment Effect and Evidence of Nonreversible Indifference Curves. *The American Economic Review*, *79*(5), 1277–1284.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In *Psychology of learning and motivation* (Vol. 54, pp. 59–101). Elsevier.

Knoblich, G., & Kircher, T. T. J. (2004). Deceiving Oneself About Being in Control: Conscious Detection of Changes in Visuomotor Coupling. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 657–666. https://doi.org/10.1037/0096-1523.30.4.657

Knoblich, G., & Sebanz, N. (2005). Agency in the face of error. *Trends in Cognitive Sciences*, *9*(6), 259–261.

Knoblich, G., & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, *15*(3), 99–104.

Koole, S. L., Dijksterhuis, A., & Van Knippenberg, A. (2001). What's in a name: Implicit self-esteem and the automatic self. *Journal of Personality and Social Psychology*, *80*(4), 669.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility--a model and taxonomy. *Psychological Review*, *97*(2), 253.

Krabbe, D., Ellbin, S., Nilsson, M., Jonsdottir, I. H., & Samuelsson, H. (2017). Executive function and attention in patients with stress-related exhaustion: perceived fatigue and effect of distraction. *Stress*, *20*(4), 333–340. https://doi.org/10.1080/10253890.2017.1336533

Krueger, J. (1998). Enhancement bias in descriptions of self and others. *Personality and Social Psychology Bulletin*, *24*(5), 505–516.

Kuiper, N. A., & MacDonald, M. R. (1982). Self and Other Perception in Mild Depressives. *Social Cognition*, *1*(3), 223–239. https://doi.org/10.1521/soco.1982.1.3.223

Kwan, V. S., Barrios, V., Ganis, G., Gorman, J., Lange, C., Kumar, M., … Keenan, J. P. (2007). Assessing the neural correlates of self-enhancement bias: a transcranial magnetic stimulation study. *Experimental Brain Research*, *182*(3), 379–385.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*(4), 880–913. https://doi.org/10.1037/0033-295X.111.4.880

Lamm, C., Decety, J., & Singer, T. (2011). Meta-analytic evidence for common and distinct neural networks associated with directly experienced pain and empathy for pain. *Neuroimage*, *54*(3), 2492–2502.

Lamm, C., & Majdandžić, J. (2015). The role of shared neural activations, mirror neurons, and morality in empathy–a critical comment. *Neuroscience Research*, *90*, 15–24.

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*(3), 137–159. https://doi.org/10.1016/S0165-0327(00)00343-8

Lehmann, A., Bahçesular, K., Brockmann, E.-M., Biederbick, S.-E., Dziobek, I., Gallinat, J., & Montag, C. (2014). Subjective experience of emotions and emotional empathy in paranoid schizophrenia. *Psychiatry Research*, *220*(3), 825–833.

Leppänen, J. M., Tenhunen, M., & Hietanen, J. K. (2003). Faster Choice-Reaction Times to Positive than to Negative Facial Expressions. *Journal of Psychophysiology*, *17*(3), 113–123. https://doi.org/10.1027//0269-8803.17.3.113

Leshikar, E. D., Dulas, M. R., & Duarte, A. (2015). Self-referencing enhances recollection in both young and older adults. *Aging, Neuropsychology, and Cognition*, *22*(4), 388–412. https://doi.org/10.1080/13825585.2014.957150

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind'. *Trends in Cognitive Sciences*, *8*(12), 528–533.

Liepelt, R., Strobach, T., Frensch, P., & Schubert, T. (2011). Improved intertask coordination after extensive dual-task practice. *Quarterly Journal of Experimental Psychology*, *64*(7), 1251–1272.

Liu, M., He, X., Rotsthein, P., & Sui, J. (2016). Dynamically orienting your own face facilitates the automatic attraction of attention. *Cognitive Neuroscience*, *7*(1–4), 37–44. https://doi.org/10.1080/17588928.2015.1044428

Liu, M., Liu, H., Zhu, Y., Wang, R., Rotshtein, P., & Sui, J. (2015). Self-related information interfere with task performances: a cross-cultural investigation. *Culture and Brain*, *3*(2), 112–121. https://doi.org/10.1007/s40167-015-0030-3

Lombardo, M. V., Barnes, J. L., Wheelwright, S. J., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PloS One*, *2*(9), e883.

Lombardo, M. V., Chakrabarti, B., Bullmore, E. T., Baron-Cohen, S., & Consortium, M. A. (2011). Specialization of right temporo-parietal junction for mentalizing and its relation to social impairments in autism. *Neuroimage*, *56*(3), 1832–1838.

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face recognition. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(3), 619–633. https://doi.org/10.1037/a0015797

Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. Psychology press.

Maddieson, I., & Emmorey, K. (1985). Relationship between Semivowels and Vowels: Cross-Linguistic Investigations of Acoustic Difference and Coarticulation. *Phonetica*, *42*(4), 163–174. https://doi.org/10.1159/000261748

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Mars, R. B., Sallet, J., Schuffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. S. (2012). Connectivity-Based Subdivisions of the Human Right 'Temporoparietal Junction Area': Evidence for Different Areas Participating in Different Cortical Networks. *Cerebral Cortex*, *22*(8), 1894–1903. https://doi.org/10.1093/cercor/bhr268

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social Connection Through Joint Action and Interpersonal Coordination. *Topics in Cognitive Science*, *1*(2), 320–339. https://doi.org/10.1111/j.1756-8765.2009.01022.x

Mattan, B. D., Quinn, K. A., Apperly, I. A., Sui, J., & Rotshtein, P. (2015). Is it always me first? Effects of self-tagging on third-person perspective-taking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1100–1117. https://doi.org/10.1037/xlm0000078

Mayer, K. M., Vuong, Q. C., & Thornton, I. M. (2017). Humans are Detected More Efficiently than Machines in the Context of Natural Scenes. *Japanese Psychological Research*, *59*(2), 178–187. https://doi.org/10.1111/jpr.12145

Meltzoff, A. N. (1990). Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling, and self practice in infancy.

Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1431), 491–500.

Mengotti, P., Ticini, L. F., Waszak, F., Schütz-Bosbach, S., & Rumiati, R. I. (2013). Imitating others' actions: transcranial magnetic stimulation of the parietal opercula reveals the processes underlying automatic imitation. *The European Journal of Neuroscience*, *37*(2), 316–322. https://doi.org/10.1111/ejn.12019

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, *214*(5–6), 655–667.

Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is There a Universal Positivity Bias in Attributions? A Meta-Analytic Review of Individual, Developmental, and Cultural Differences in the Self-Serving Attributional Bias. *Psychological Bulletin*, *130*(5), 711–747. https://doi.org/10.1037/0033-2909.130.5.711

Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, *14*(2), 247–279.

Miller, L. M., & Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Memory & Cognition*, *37*(6), 850–865.

Milward, S. J., Kita, S., & Apperly, I. A. (2014). The development of co-representation effects in a joint task: Do children represent a co-actor? *Cognition*, *132*(3), 269–279. https://doi.org/10.1016/j.cognition.2014.04.008

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*(8), 1306–1315. https://doi.org/10.1162/0898929055002418

Montag, C., Heinz, A., Kunz, D., & Gallinat, J. (2007). Self-reported empathic abilities in schizophrenia. *Schizophrenia Research*, *92*(1), 85–89.

Montemayor, R., & Eisen, M. (1977). The development of self-conceptions from childhood to adolescence. *Developmental Psychology*, *13*(4), 314.

Moradi, Z., Sui, J., Hewstone, M., & Humphreys, G. W. (2015). In-group modulation of perceptual matching. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-014-0798-8

Moran, J. M., Macrae, C. N., Heatherton, T. F., Wyland, C. L., & Kelley, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *Journal of Cognitive Neuroscience*, *18*(9), 1586–1594.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*(1), 56–60. https://doi.org/10.1080/17470215908416289

Müller, U., Zelazo, P. D., & Imrisek, S. (2005). Executive function and children's understanding of false belief: How specific is the relation? *Cognitive Development*, *20*(2), 173–189.

Nelson, D. L., & Schreiber, T. A. (1992). Word concreteness and word structure as independent determinants of recall. *Journal of Memory and Language*, *31*(2), 237–260. https://doi.org/10.1016/0749-596X(92)90013-N

Neumann, D., Spezio, M. L., Piven, J., & Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Social Cognitive and Affective Neuroscience*, *1*(3), 194–202.

Northoff, G., & Hayes, D. J. (2011). Is our self nothing but reward? *Biological Psychiatry*, *69*(11), 1019–1025. https://doi.org/10.1016/j.biopsych.2010.12.014

Northoff, G., Heinzel, A., De Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *Neuroimage*, *31*(1), 440–457.

Northoff, G., Schneider, F., Rotte, M., Matthiae, C., Tempelmann, C., Wiebking, C., … Heinze, H.-J. (2009). Differential parametric modulation of self-relatedness and emotions in different brain regions. *Human Brain Mapping*, *30*(2), 369–382.

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2012). Distinguishing Self and Other in Joint Action. Evidence from a Musical Paradigm. *Cerebral Cortex*, *22*(12), 2894–2903. https://doi.org/10.1093/cercor/bhr364

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology*, *80*(3), 381–396. https://doi.org/10.1037//0022-3514.80.3.381

Olivers, C. N. L., & Nieuwenhuis, S. (2006). The beneficial effects of additional task load, positive affect, and instruction on the attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 364–379. https://doi.org/10.1037/0096-1523.32.2.364

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776.

Oztop, E., Kawato, M., & Arbib, M. (2006). Mirror neurons and imitation: A computationally guided review. *Neural Networks*, *19*(3), 254–271.

Pahl, S., & Eiser, J. R. (2005). Valence, comparison focus and self-positivity biases: does it matter whether people judge positive or negative traits? *Experimental Psychology*, *52*(4), 303.

Paivio, A. (1967). Paired-Associate Learning and Free Recall of Nouns as a Function of Concreteness, Specificity, Imagery, and Meaningfulness. *Psychological Reports*, *20*(1), 239–245. https://doi.org/10.2466/pr0.1967.20.1.239

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255–287.

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1196–1204. https://doi.org/10.1037/0278-7393.20.5.1196

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, *76*(1p2), 1.

Payne, S., Tsakiris, M., & Maister, L. (2017). Can the self become another? Investigating the effects of self-association with a new facial identity. *The Quarterly Journal of Experimental Psychology*, *70*(6), 1085–1097. https://doi.org/10.1080/17470218.2015.1137329

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *Journal of Neuroscience*, *29*(36), 11182–11191.

Perner, J. (1999). Theory of mind. *Developmental Psychology: Achievements and Prospects*, 205–230.

Philipp, A. M., & Prinz, W. (2010). Evidence for a role of the responding agent in the joint compatibility effect. *The Quarterly Journal of Experimental Psychology*, *63*(11), 2159–2171. https://doi.org/10.1080/17470211003802426

Preckel, K., Kanske, P., & Singer, T. (2018). On the interaction of social affect and cognition: empathy, compassion and theory of mind. *Current Opinion in Behavioral Sciences*, *19*, 1–6. https://doi.org/10.1016/j.cobeha.2017.07.010

Prikken, M., van der Weiden, A., Kahn, R. S., Aarts, H., & van Haren, N. E. M. (2018). Impaired self-agency inferences in schizophrenia: The role of cognitive capacity and causal reasoning style. *European Psychiatry: The Journal of the Association of European Psychiatrists*, *47*, 27–34. https://doi.org/10.1016/j.eurpsy.2017.08.007

Prinz, W. (1990). A common coding approach to perception and action. In *Relationships between perception and action* (pp. 167–201). Springer.

Prinz, W. (2015). Task representation in individual and joint settings. *Frontiers in Human Neuroscience*, *9*. https://doi.org/10.3389/fnhum.2015.00268

Prizant, B. M., & Duchan, J. F. (1981). The functions of immediate echolalia in autistic children. *Journal of Speech and Hearing Disorders*, *46*(3), 241–249.

Przyrembel, M., Smallwood, J., Pauen, M., & Singer, T. (2012). Illuminating the dark matter of social neuroscience: Considering the problem of social interaction

from philosophical, psychological, and neuroscientific perspectives. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00190

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010a). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230–236.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010b). Executive function is necessary for perspective selection, not Level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition*, *117*(2), 230–236.

Raab, D. H. (1962). Division of psychology: Statistical facilitation of simple reaction times. *Transactions of the New York Academy of Sciences*, *24*(5 Series II), 574–590.

Rakoczy, H. (2010). Executive function and the development of belief–desire psychology. *Developmental Science*, *13*(4), 648–661.

Ramachandran, V. S., & Oberman, L. M. (2006). Broken mirrors: a theory of autism. *Scientific American*, *295*(5), 62–69.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 849–860. https://doi.org/10.1037/0096-1523.18.3.849

Rehfeldt, R. A., Dillen, J. E., Ziomek, M. M., & Kowalchuk, R. K. (2007). Assessing relational learning deficits in perspective-taking in children with high-functioning autism spectrum disorder. *The Psychological Record*, *57*(1), 23–47.

Reuther, J., & Chakravarthi, R. (2017). Does self-prioritization affect perceptual processes? *Visual Cognition*, *25*(1–3), 381–398. https://doi.org/10.1080/13506285.2017.1323813

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: interpretations and misinterpretations. *Nature Reviews. Neuroscience*, *11*(4), 264–274. https://doi.org/10.1038/nrn2805

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, *35*(9), 677–688. http://0-dx.doi.org.pugwash.lib.warwick.ac.uk/10.1037/0022-3514.35.9.677

Royzman, E. B., Cassidy, K. W., & Baron, J. (2003). ' I know, you know': Epistemic egocentrism in children and adults. *Review of General Psychology*, *7*(1), 38.

Ruys, K. I., & Aarts, H. (2010). When competition merges people's behavior: Interdependency activates shared action representations. *Journal of Experimental Social Psychology*, *46*(6), 1130–1133. https://doi.org/10.1016/j.jesp.2010.05.016

Salomon, R., Lim, M., Kannape, O., Llobera, J., & Blanke, O. (2013). 'Self pop-out': agency enhances self-recognition in visual search. *Experimental Brain Research*, *228*(2), 173–181. https://doi.org/10.1007/s00221-013-3549-6

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: a case of a selective deficit in inhibiting self-perspective. *Brain*, *128*(5), 1102–1111.

Samson, D., Houthuys, S., & Humphreys, G. W. (2015). Self-perspective inhibition deficits cannot be explained by general executive control difficulties. *Cortex*, *70*, 189–201.

Sanders, G. S. (1981). Driven by distraction: An integrative review of social facilitation theory and research. *Journal of Experimental Social Psychology*, *17*(3), 227–251.

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2012). Enhancing social ability by stimulating right temporoparietal junction. *Current Biology*, *22*(23), 2274–2277.

Santiesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: From imitation to Theory of Mind. *Cognition*, *122*(2), 228–235. https://doi.org/10.1016/j.cognition.2011.11.004

Sato, A. (2009). Both motor prediction and conceptual congruency between preview and action-effect contribute to explicit judgment of agency. *Cognition*, *110*(1), 74–83. https://doi.org/10.1016/j.cognition.2008.10.011

Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, *94*(3), 241–255. https://doi.org/10.1016/j.cognition.2004.04.003

Saunders, D. R., Melcher, D., & van Zoest, W. (2017). No evidence of task co-representation in a joint Stroop task. *Psychological Research*. https://doi.org/10.1007/s00426-017-0909-z

Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2012). The presence of threat affects saccade trajectories. *Visual Cognition*, *20*(3), 284–299. https://doi.org/10.1080/13506285.2012.658885

Schmidt, L. J., Belopolsky, A. V., & Theeuwes, J. (2015). Attentional capture by signals of threat. *Cognition & Emotion*, *29*(4), 687–694. https://doi.org/10.1080/02699931.2014.924484

Schmidt, R. C., Fitzpatrick, P., Caron, R., & Mergeche, J. (2011). Understanding social motor coordination. *Human Movement Science*, *30*(5), 834–845.

Schulte-Rüther, M., Fink, G. R., Pohl, A., Herpertz-Dahlmann, B., Konrad, K., & Harindranathan, P. (2015). Effective Connectivity of Mirror System Brain Areas in Autism Spectrum Disorder. In *International Meeting for Autism Research (Annual Meeting of the International Society for Autism Research)*. Kognitive Neurowissenschaften.

Schulte-Rüther, M., Otte, E., Adigüzel, K., Firk, C., Herpertz-Dahlmann, B., Koch, I., & Konrad, K. (2017). Intact mirror mechanisms for automatic facial emotions in children and adolescents with autism spectrum disorder. *Autism Research*, *10*(2), 298–310.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*(5), 499–520.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006a). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Sebanz, N., Bekkering, H., & Knoblich, G. (2006b). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Sebanz, N., & Frith, C. D. (2004). Beyond simulation? Neural mechanisms for predicting the actions of others. *Nature Neuroscience*, *7*(1), 5–6. https://doi.org/10.1038/nn0104-5

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: just like one's own? *Cognition*, *88*(3), B11–B21. https://doi.org/10.1016/S0010-0277(03)00043-X

Sebanz, N., Knoblich, G., & Prinz, W. (2005). How Two Share a Task: Corepresenting Stimulus-Response Mappings. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1234–1246. https://doi.org/10.1037/0096-1523.31.6.1234

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (2006). Twin peaks: An ERP study of action planning and control in coacting individuals. *Journal of Cognitive Neuroscience*, *18*(5), 859–870.

Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science*, *3*(2), 102–116.

Sellaro, R., Treccani, B., Rubichi, S., & Cubelli, R. (2013). When co-action eliminates the Simon effect: disentangling the impact of co-actor's presence and task sharing on joint-task performance. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00844

Sergent, C., Baillet, S., & Dehaene, S. (2005). *Timing of the brain events underlying access to consciousness during the attentional blink* (Vol. 8). https://doi.org/10.1038/nn1549

Shapiro, K. L., Caldwell, J., & Sorensen, R. E. (1997). Personal names and the attentional blink: a visual 'cocktail party' effect. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(2), 504–514.

Silani, G., Lamm, C., Ruff, C. C., & Singer, T. (2013). Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. *Journal of Neuroscience*, *33*(39), 15466–15476.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In *Advances in psychology* (Vol. 65, pp. 31–86). Elsevier.

Simon, J. R., & Rudell, A. P. (1967). Auditory SR compatibility: the effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, *51*(3), 300.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, *24*(18), R875–R878.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for Pain Involves the Affective but not Sensory Components of Pain. *Science*, *303*(5661), 1157–1162. https://doi.org/10.1126/science.1093535

Southgate, V., & Hamilton, A. F. de C. (2008). Unbroken mirrors: Challenging a theory of autism. *Trends in Cognitive Sciences*, *12*(6), 225–229.

Sowden, S., Koehne, S., Catmur, C., Dziobek, I., & Bird, G. (2016). Intact automatic imitation and typical spatial compatibility in autism spectrum disorder: challenging the broken mirror theory. *Autism Research*, *9*(2), 292–300.

Sowden, S., & Shah, P. (2014). Self-other control: a candidate mechanism for social cognitive function. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00789

Sparks, S., Cunningham, S. J., & Kritikos, A. (2016). Culture modulates implicit ownership-induced self-bias in memory. *Cognition*, *153*, 89–98. https://doi.org/10.1016/j.cognition.2016.05.003

Spengler, S., Bird, G., & Brass, M. (2010). Hyperimitation of actions is related to reduced understanding of others' minds in autism spectrum conditions. *Biological Psychiatry*, *68*(12), 1148–1155.

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Control of shared representations relies on key processes involved in mental state attribution. *Human Brain Mapping*, *30*(11), 3704–3718.

Spengler, S., von Cramon, D. Y., & Brass, M. (2010). Resisting motor mimicry: control of imitation involves processes central to social cognition in patients with frontal and temporo-parietal lesions. *Social Neuroscience*, *5*(4), 401–416.

Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self-and external-agency attribution: a brief review and meta-analysis. *Brain Structure and Function*, *216*(2), 151–157.

Steinbeis, N. (2016). The role of self–other distinction in understanding others' mental and emotional states: neurocognitive mechanisms in children and adults. *Phil. Trans. R. Soc. B*, *371*(1686), 20150074.

Stenzel, A., Dolk, T., Colzato, L. S., Sellaro, R., Hommel, B., & Liepelt, R. (2014). The joint Simon effect depends on perceived agency, but not intentionality, of the alternative action. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00595

Stolte, M., Humphreys, G. W., Yankouskaya, A., & Sui, J. (2017). Dissociating biases towards the self and positive emotion. *The Quarterly Journal of Experimental Psychology*, *70*(6), 1011–1022. https://doi.org/10.1080/17470218.2015.1101477

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643.

Stürmer, B., Aschersleben, G., & Prinz, W. (2000). Correspondence effects with manual gestures and postures: a study of imitation. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(6), 1746.

Sui, J. (2016). Self-Reference Acts as a Golden Thread in Binding. *Trends in Cognitive Sciences*, *20*(7), 482–483. https://doi.org/10.1016/j.tics.2016.04.005

Sui, J., Enock, F., Ralph, J., & Humphreys, G. W. (2015). Dissociating hyper and hypoself biases to a core self-representation. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *70*, 202–212. https://doi.org/10.1016/j.cortex.2015.04.024

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(5), 1105–1117. https://doi.org/10.1037/a0029792

Sui, J., Hong, Y., Hong Liu, C., Humphreys, G. W., & Han, S. (2013). Dynamic cultural modulation of neural responses to one's own and friend's faces. *Social Cognitive and Affective Neuroscience*, *8*(3), 326–332. https://doi.org/10.1093/scan/nss001

Sui, J., & Humphreys, G. W. (2013a). Self-referential processing is distinct from semantic elaboration: Evidence from long-term memory effects in a patient with amnesia and semantic impairments. *Neuropsychologia*, *51*(13), 2663–2673. https://doi.org/10.1016/j.neuropsychologia.2013.07.025

Sui, J., & Humphreys, G. W. (2013b). The boundaries of self face perception: Response time distributions, perceptual categories, and decision weighting. *Visual Cognition*, *21*(4), 415–445. https://doi.org/10.1080/13506285.2013.800621

Sui, J., & Humphreys, G. W. (2015a). More of me! Distinguishing self and reward bias using redundancy gains. *Attention, Perception, & Psychophysics*, *77*(8), 2549–2561. https://doi.org/10.3758/s13414-015-0970-x

Sui, J., & Humphreys, G. W. (2015b). Super-size me: self biases increase to larger stimuli. *Psychonomic Bulletin & Review*, *22*(2), 550–558. https://doi.org/10.3758/s13423-014-0690-6

Sui, J., & Humphreys, G. W. (2015c). The integrative self: how self-reference integrates perception and memory. *Trends in Cognitive Sciences*, *19*(12), 719–728.

Sui, J., & Humphreys, G. W. (2015d). The interaction between self-bias and reward: Evidence for common and distinct processes. *The Quarterly Journal of Experimental Psychology*, *68*(10), 1952–1964. https://doi.org/10.1080/17470218.2015.1023207

Sui, J., Liu, C. H., Wang, L., & Han, S. (2009). Attentional orientation induced by temporarily established self-referential cues. *The Quarterly Journal of Experimental Psychology*, *62*(5), 844–849. https://doi.org/10.1080/17470210802559393

Sui, J., Liu, H., & Han, S. (2009). Cultural difference in neural mechanisms of self-recognition. *Social Neuroscience*, *4*(5), 402–411. https://doi.org/10.1080/17470910802674825

Sui, J., Rotshtein, P., & Humphreys, G. W. (2013). Coupling social attention to the self forms a network for personal significance. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(19), 7607–7612. https://doi.org/10.1073/pnas.1221862110

Sui, J., Sun, Y., Peng, K., & Humphreys, G. W. (2014). The automatic and the expected self: separating self- and familiarity biases effects by manipulating stimulus probability. *Attention, Perception, & Psychophysics*, *76*(4), 1176–1184. https://doi.org/10.3758/s13414-014-0631-5

Sui, J., Yankouskaya, A., & Humphreys, G. W. (2015). Super-capacity me! Super-capacity and violations of race independence for self- but not for reward-associated stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(2), 441–452. https://doi.org/10.1037/a0038288

Sui, J., & Zhu, Y. (2005). Five-year-olds can show the self-reference advantage. *International Journal of Behavioral Development*, *29*(5), 382–387. https://doi.org/10.1177/01650250500172673

Sui, J., Zhu, Y., & Han, S. (2006). Self-face recognition in attended and unattended conditions: an event-related brain potential study. *Neuroreport*, *17*(4), 423–427.

Surtees, A. D. R., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, *83*(2), 452–460.

Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin*, *121*(3), 371–394. https://doi.org/10.1037/0033-2909.121.3.371

Tacikowski, P., Cygan, H. B., & Nowicka, A. (2014). Neural correlates of own and close-other's name recognition: ERP evidence. *Frontiers in Human Neuroscience*, *8*, 194.

Tacikowski, P., & Nowicka, A. (2010). Allocation of attention to self-name and self-face: An ERP study. *Biological Psychology*, *84*(2), 318–324.

Tipples, J., Atkinson, A. P., & Young, A. W. (2002). The eyebrow frown: a salient social signal. *Emotion (Washington, D.C.)*, *2*(3), 288–296.

Tomasello, M. (1995). Joint attention as social cognition. *Joint Attention: Its Origins and Role in Development*, *103130*.

Tong, F., & Nakayama, K. (1999). Robust representations for faces: evidence from visual search. *Journal of Experimental Psychology. Human Perception and Performance*, *25*(4), 1016–1035.

Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, *12*(4), 242–248. https://doi.org/10.1080/17470216008416732

Treisman, A. (1964a). Monitoring and storage of irrelevant messages in selective attention. *Journal of Verbal Learning and Verbal Behavior*, *3*(6), 449–459. https://doi.org/10.1016/S0022-5371(64)80015-3

Treisman, A. (1964b). Selective attention in man. *British Medical Bulletin*.

Triandis, H. C., Bontempo, R., Villareal, M. J., Asai, M., & Lucca, N. (1988). Individualism and collectivism: Cross-cultural perspectives on self-ingroup relationships. *Journal of Personality and Social Psychology*, *54*(2), 323.

Tusche, A., Böckler, A., Kanske, P., Trautwein, F.-M., & Singer, T. (2016). Decoding the charitable brain: empathy, perspective taking, and attention shifts differentially predict altruistic giving. *Journal of Neuroscience*, *36*(17), 4719–4732.

van der Wel, R. P. R. D. (2015). Me and we: Metacognition and performance evaluation of joint actions. *Cognition*, *140*, 49–59. https://doi.org/10.1016/j.cognition.2015.03.011

van Schie, H. T., Mars, R. B., Coles, M. G., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*(5), 549.

van Ulzen, N. R., Lamoth, C. J., Daffertshofer, A., Semin, G. R., & Beek, P. J. (2008). Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side. *Neuroscience Letters*, *432*(2), 88–93.

Vanvuchelen, M., Roeyers, H., & De Weerdt, W. (2011). Do imitation problems reflect a core characteristic in autism? Evidence from a literature review. *Research in Autism Spectrum Disorders*, *5*(1), 89–95.

Vesper, C., Abramova, E., Bütepage, J., Ciardo, F., Crossey, B., Effenberg, A., … Wahn, B. (2017). Joint Action: Mental Representations, Shared Information and General Mechanisms for Coordinating with Others. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.02039

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, *23*(8–9), 998–1003.

Vincini, S., Jhang, Y., Buder, E. H., & Gallagher, S. (2017). Neonatal imitation: Theory, experimental design, and significance for the field of social cognition. *Frontiers in Psychology*, *8*, 1323.

Vlainic, E., Liepelt, R., Colzato, L. S., Prinz, W., & Hommel, B. (2010). The Virtual Co-Actor: The Social Simon Effect does not Rely on Online Feedback from the Other. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00208

Vuilleumier, P., & Pourtois, G. (2007). Distributed and interactive brain mechanisms during emotion face perception: evidence from functional neuroimaging. *Neuropsychologia*, *45*(1), 174–194.

Wade, G. L., & Vickery, T. J. (2017). Self-relevance effects and label choice: Strong variations in label-matching performance due to non-self-relevant factors. *Attention, Perception, & Psychophysics*, 1–11. https://doi.org/10.3758/s13414-017-1307-8

Wang, L., Liu, X., Guise, K. G., Knight, R. T., Ghajar, J., & Fan, J. (2010). Effective connectivity of the fronto-parietal network during attentional control. *Journal of Cognitive Neuroscience*, *22*(3), 543–553.

Watson, D. G., Blagrove, E., Evans, C., & Moore, L. (2012). Negative triangles: simple geometric shapes convey emotional valence. *Emotion*, *12*(1), 18.

Watson, L.A., Dritschel, B., Obonsawin, M. C., & Jentzsch, I. (2007). Seeing yourself in a positive light: Brain correlates of the self-positivity bias. *Brain Research*, *1152*, 106–110. https://doi.org/10.1016/j.brainres.2007.03.049

Watson, Lynn A., Dritschel, B., Obonsawin, M. C., & Jentzsch, I. (2007). Seeing yourself in a positive light: brain correlates of the self-positivity bias. *Brain Research*, *1152*, 106–110.

Wegner, D. M., & Sparrow, B. (2007). The puzzle of coaction. *Distributed Cognition and the Will: Individual Volition in Social Context*, 17–37.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: experiencing control over the movements of others. *Journal of Personality and Social Psychology*, *86*(6), 838.

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*(7), 480.

Wel, R. P. R. D. van der, & Fu, E. (2015). Entrainment and task co-representation effects for discrete and continuous action sequences. *Psychonomic Bulletin & Review*, 1–7. https://doi.org/10.3758/s13423-015-0831-6

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development*, *72*(3), 655–684.

Wenger, M. J., & Townsend, J. T. (2006). On the costs and benefits of faces and words: Process characteristics of feature search in highly meaningful stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 755.

Wenke, D., Atmaca, S., Holländer, A., Liepelt, R., Baess, P., & Prinz, W. (2011). What is Shared in Joint Action? Issues of Co-representation, Response Conflict, and Agent Identification. *Review of Philosophy and Psychology*, *2*(2), 147–172. https://doi.org/10.1007/s13164-011-0057-0

Williams, D. (2010). Theory of own mind in autism: Evidence of a specific deficit in self-awareness? *Autism*, *14*(5), 474–494.

Winter, K., Spengler, S., Bermpohl, F., Singer, T., & Kanske, P. (2017). Social cognition in aggressive offenders: Impaired empathy, but intact theory of mind. *Scientific Reports*, *7*(1), 670.

Wong, A. C.-N., & Wong, Y. K. (2014). Interaction between perceptual and cognitive processing well acknowledged in perceptual expertise research. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00308

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 255–260. https://doi.org/10.1037/0278-7393.21.1.255

Yankouskaya, A., Booth, D. A., & Humphreys, G. W. (2012). Interactions between facial emotion and identity in face processing: Evidence based on redundancy gains. *Attention, Perception, & Psychophysics*, *74*(8), 1692–1711.

Zhu, Y., & Han, S. (2008). Cultural differences in the self: From philosophy to psychology and neuroscience. *Social and Personality Psychology Compass*, *2*(5), 1799–1811.

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *NeuroImage*, *34*(3), 1310–1316. https://doi.org/10.1016/j.neuroimage.2006.08.047

Zwickel, J., White, S. J., Coniston, D., Senju, A., & Frith, U. (2010). Exploring the building blocks of social cognition: spontaneous agency perception and visual

perspective taking in autism. *Social Cognitive and Affective Neuroscience*, *6*(5), 564–571.